

UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À  
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE  
DE LA MAÎTRISE EN BIOPHYSIQUE ET BIOLOGIE CELLULAIRES

PAR  
MAXIME SASSEVILLE

STABILISATION DE LA PROTÉINE MB-1TRP PAR LE DESIGN DE PONTS  
SALINS À SA SURFACE

9 AVRIL 2004

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

Trois-Rivières, le 9 avril 2004

Objet : Contribution à l'article utilisé pour le mémoire de Maxime Sasseville soumis comme exigence partielle à l'obtention de la maîtrise en biophysique et biologie cellulaires.

Bonjour,

Par la présente, j'aimerais préciser ma contribution, ainsi que celle des autres auteurs, aux résultats de l'article scientifique que j'utilise dans mon mémoire.

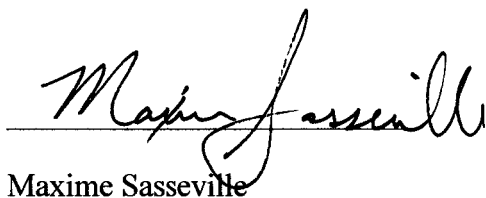
**Article :** Single-point mutations at the surface of MB-1Trp lead to important changes in conformational properties. *Maxime Sasseville, Sébastien Claveau, Mylène C. Gagnon, Marc Beauregard.* Sous presse dans Journal of Peptide Science.

**Contribution personnelle :** J'ai réalisé la totalité des expériences de biologie moléculaire. Les expériences de purification et de caractérisation de l'article ont été faites par moi-même. L'analyse et l'interprétation des résultats ont été faites essentiellement par moi-même en collaboration avec mon directeur de recherche. La rédaction de l'article s'est également faite avec la proche collaboration de mon directeur de recherche, spécialement pour l'élaboration des idées maîtresses, du plan et pour la révision éditoriale du texte. J'ai effectué la rédaction proprement dite du manuscrit pour le premier jet.

**Contribution de Sébastien Claveau, B.Sc. :** Sébastien a contribué aux vagues de purification et de caractérisation subséquentes. Son aide a été également essentielle lors de discussions sur les résultats et lors de la rédaction.

**Contribution de Mylène C. Gagnon, M.Sc. MBA :** Mylène a contribué à la rédaction de l'article par la révision du manuscrit ainsi que par des discussions éditoriales.

**Contribution de Marc Beauregard, Ph.D. :** Monsieur Beauregard a élaboré le projet dans un premier temps et a été mon guide pendant mes études à l'université de Québec à Trois-Rivières. Au niveau de la rédaction de l'article, il a apporté un regard critique aux résultats ainsi qu'aux interprétations. Il a également complété et finalisé le manuscrit.



Maxime Sasseville

Étudiant à la maîtrise en Biophysique et Biologie Cellulaire sous la direction du professeur Marc Beauregard

Je confirme que les affirmations du présent document sont vraies.

---

Marc Beauregard, Ph.D

Professeur au département de Chimie-Biologie, Université du Québec à Trois-Rivières

## RÉSUMÉ

La protéine MB-1 est une protéine à haute valeur nutritive destinée à l'alimentation animale. Cette protéine est également utilisée d'un point de vue fondamental pour étudier les lois qui relient la séquence primaire à la structure de la protéine. Le variant le plus stable de MB-1 jusqu'à présent est MB-1Trp, décrit précédemment (Gagnon et al. 2000). Pour permettre son application dans le domaine de l'agroalimentaire, MB-1Trp doit être d'une stabilité comparable et même supérieure à celle d'une protéine naturelle. Cette étude présente l'application d'une stratégie impliquant l'ajout d'interactions électrostatiques à la surface de la protéine MB-1Trp afin d'en augmenter la stabilité thermique et protéolytique.

Cette stratégie comprend dans un premier temps le retrait de la surface de la protéine de deux résidus, la leucine 19 et la leucine 68, dont la chaîne latérale hydrophobe est exposée au solvant. Dans un deuxième temps, on remplace ces résidus hydrophobes par des résidus chargés négativement, des glutamates, susceptibles de faire des ponts salins intrahélicaux. Le but du projet était donc de procéder aux constructions génétiques, à l'expression des protéines et à leur caractérisation.

L'ajout de ponts salins, dans le cas du mutant Leu68Glu, a augmenté la température de dénaturation de la protéine de 6°C et dans le cas du mutant Leu19Glu, l'a diminué de 7,5°C. Les expériences de caractérisation nous ont démontré que ces protéines ont un haut contenu en structures secondaires hélicales. Des expériences de

masquage des interactions électrostatiques par l'ajout de sel nous ont montré que la formation de ponts salins avait lieu seulement dans le variant Leu68Glu. Le cœur hydrophobe de ce dernier a démontré sensiblement les mêmes caractéristiques que celui de MB-1Trp, tandis que le Leu19Glu a démontré une réorganisation du cœur qui expose le tryptophane tel que montré par fluorescence.

Le mutant Leu68Glu dépasse son prédécesseur MB-1Trp en stabilité. Les informations recueillies par fluorescence, dichroïsme circulaire, dénaturation thermique et dégradation protéolytique nous indiquent que ce mutant pourrait se comporter mieux que ses prédécesseurs lorsque utilisé comme protéine-puits pour les acides aminés essentiels dans un organisme transgénique à intérêt agroalimentaire.

Des raisons qui différencient l'effet des substitutions faites en position 19 et 68 sont proposées et discutées. Il appert que la position de la leucine 19 dans une hélice N-terminale soit un facteur à considérer. Des avenues de recherche sont proposées pour une meilleure compréhension de la mutation en position 19, ainsi que pour des améliorations futures aux protéines de la famille MB-1.

## REMERCIEMENTS

Ces deux années d'étude se terminent et il est maintenant l'heure des bilans. Cette période a été une fascinante et enrichissante démarche dans le vaste domaine de l'ingénierie et l'étude des protéines. La première personne que je souhaite remercier est Monsieur Marc Beauregard. C'est grâce à son soutien, à son dynamisme et à sa dévorante passion pour son travail qu'il a rendu ce séjour si agréable et fructueux. Il m'a transmis cette débordante passion et elle me poussera sous peu vers la poursuite de mes études supérieures. Je le remercie surtout pour la confiance qu'il a placée en moi depuis le tout début. Elle m'a permis d'évoluer d'une façon accélérée (si vous me permettez le jeu de mot).

Je ne peux passer sous silence la collaboration de tous les membres de l'équipe de recherche de M.Beauregard, sans qui l'accomplissement de ce travail n'aurait sans doute pas été possible. Merci à Sébastien Claveau, mon compagnon d'évolution accélérée. Merci également à Catherine St-Louis, ma compagne de protéolyse et du projet NL et à Alain Doucet, mon compagnon de purification de protéine. De plus, je remercie François-Olivier McDuff pour ces trépidantes discussions à la limite de la science et de la philosophie. Finalement, merci à mes nouveaux collaborateurs, Habib Khoudi et Mylène Gagnon ainsi qu'aux stagiaires que j'ai eu le plaisir de côtoyer : Caroline C., Caroline S., Patrick, Alain, Vicky et Cindy.



## TABLE DES MATIÈRES

LETTRE DE CONTRIBUTIONS .....	II
RÉSUMÉ .....	V
REMERCIEMENTS .....	VII
TABLE DES MATIÈRES .....	VIII
LISTE DES FIGURES .....	XI
LISTE DES ABRÉVIATIONS .....	XIII
NOTE SUR LA NOMENCLATURE DES MUTANTS DANS LE PRÉSENT DOCUMENT .....	XIV
 <b>CHAPITRE I      INTRODUCTION .....</b>	<b>1</b>
1.1 Les acides aminés et la nutrition .....	1
1.2 Les solutions pour une meilleure alimentation des animaux d'élevage .....	4
1.2.1 L'ingénierie métabolique.....	5
1.2.2 L'expression hétérologue d'une protéine naturelle riche en AAE .....	8
1.2.3 L'expression hétérologue d'une protéine de design <i>de novo</i> riche en AAE .....	10
1.3 Le design de protéines .....	13
1.4 Le projet "Milk Bundle-1" .....	18
1.5 Exemple d'améliorations de MB-1 : MB-1Trp .....	20
1.6 La stratégie de cette étude .....	23

## **CHAPITRE II      RÉSULTATS .....30**

### **2.1 Article scientifique .....30**

Résumé français .....31

Summary .....33

Introduction .....35

Materials and Methods .....38

Results .....42

Discussion .....45

Acknowledgment .....48

References .....49

Tables .....57

Figure Legends .....59

Figures .....61

### **2.2 Autres résultats .....66**

## **CHAPITRE III      DISCUSSION .....68**

### **3.1 L'effet des ponts salins sur la stabilité de MB-1Trp .....68**

### **3.2 Les éléments qui différencient les mutations aux positions 19 et 68 .....73**

3.2.1 Quelles sont les contributions impliquées dans la stabilisation du mutant Leu68Glu et qui devraient être présentes dans le mutant Leu19Glu? .....	73
3.2.2 Qu'est-ce qui peut expliquer une différence entre ces deux substitutions apparemment semblables (Leu19Glu vs Leu68Glu)? .....	76
3.3 Comparaison avec les précédents homologues de MB-1 .....	80
3.4 Perspectives de recherche .....	82
<b>BIBLIOGRAPHIE .....</b>	<b>85</b>

## LISTE DES FIGURES

### CHAPITRE I INTRODUCTION

Figure 1.1 : Dogme central de la biologie moléculaire .....	2
Figure 1.2 : La voie générale de biosynthèse de la famille de l'aspartate .....	6
Figure 1.3 : La chaîne latérale de la leucine et de la méthionine .....	12
Figure 1.4 : Les motifs tridimensionnels les plus courants pour des protéines de design <i>de novo</i> .....	14
Figure 1.5 : La protéine MB-1 avec sa structure primaire et secondaire .....	19
Figure 1.6 : A) Chaîne latérale de la tyrosine et du tryptophane. B) Représentation de MB-1 avec seulement les boucles et les acides aminés participant au cœur hydrophobe.	22
Figure 1.7 : Clathrate de molécules d'eau hydratant une chaîne latérale de leucine .....	24
Figure 1.8 : A) Vue latérale d'un polypeptide en conformation hélice- $\alpha$ . B) Vue de haut, le même polypeptide où il est montré les interactions électrostatiques de type ponts salins entre trois chaînes latérales.....	26
Figure 1.9 : Comparaison des caractéristiques de la leucine, du glutamate et de l'aspartate .....	27

## CHAPITRE II L'ARTICLE SCIENTIFIQUE

Figure 2.1: Predicted location of secondary structures in MB-1 TrpHis. The structure is flattened with the core-forming hydrophobic residues in background .....	61
Figure 2.2: CD spectra of MB-1 Trp and its variants .....	62
Figure 2.3: Tryptophane fluorescence spectra of MB-1 Trp and its variants .....	63
Figure 2.4: Thermal denaturation of MB-1 Trp and its variants monitored using molar ellipticity at 222 nm .....	64
Figure 2.5: Proteolytic degradation of MB-1 Trp and its variants.....	65

## CHAPITRE III DISCUSSION

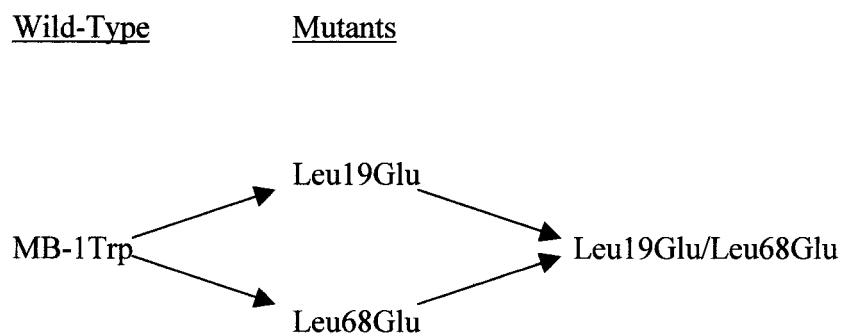
Figure 3.1 : A) Exemple de la protéine en baril à trois hélices ayant un réseau de ponts salins E-K et E-R aux interfaces (positions " g " et " e " sur l'heptade) B) La même protéine montrant la contribution de spécification du repliement des interactions électrostatiques .....	72
Figure 3.2 : La chaîne latérale du glutamate et de la glutamine .....	75
Figure 3.3 : Compilation des stratégies appliquées à MB-1 depuis sa création et des stabilités thermodynamiques correspondantes .....	81

**LISTE DES ABRÉVIATIONS**

AAE	Acides aminés essentiels
MB-1	Milk Bundle-1
MB-1Trp	MB-1 portant une substitution de la Tyr62 par un Trp
MB-1His	MB-1 portant un ajout de 6 His à son extrémité N-terminale
Leu19Glu	Substitution de la Leu19 par un Glu
Leu68Glu	Substitution de la Leu68 par un Glu
BCA	Bicinconinique acid (acide bicinconinique)
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (gel électrophorétique de polyacrylamide contenant du dodécyl sulfate de sodium)
CD	Circular dichroism (Dichroïsme circulaire)
T <sub>m</sub>	Température de dénaturation
ANSA	8-anilino-1-naphthalenesulfonic acid
RMN	Résonance magnétique nucléaire

## NOTE SUR LA NOMENCLATURE DES MUTANTS DANS LE PRÉSENT DOCUMENT

Par soucis de clarté et pour alléger les formules en utilisation usuelle dans le laboratoire, le nom des mutants sera réduit dans ce document au nom de la mutation elle-même que nous avons appliquée à la protéine MB-1 Trp. Ainsi, la protéine MB-1 Trp à laquelle nous avons fait la mutation de la leucine en position 19 pour un glutamate sera appelée Leu19Glu. Il en est ainsi pour le second mutant, Leu68Glu, et le mutant double, Leu19Glu/Leu19Glu.



Il est à noter également qu'une queue de 6 histidines a été ajoutée à chacun des mutants pour en rendre la purification possible (point qui sera discuté au chapitre III). Cette queue n'est pas présente dans la protéine de départ MB-1 Trp, mais elle est supposée n'avoir aucun effet sur la conformation de la protéine tel qu'observé lors de l'ajout de la même queue à MB-1 (communication personnelle de M. Beauregard; Grundy et al. 1998). Une fois encore dans un esprit de clarté et de légèreté, aucune addition au nom n'a été faite à cet égard.

## CHAPITRE I

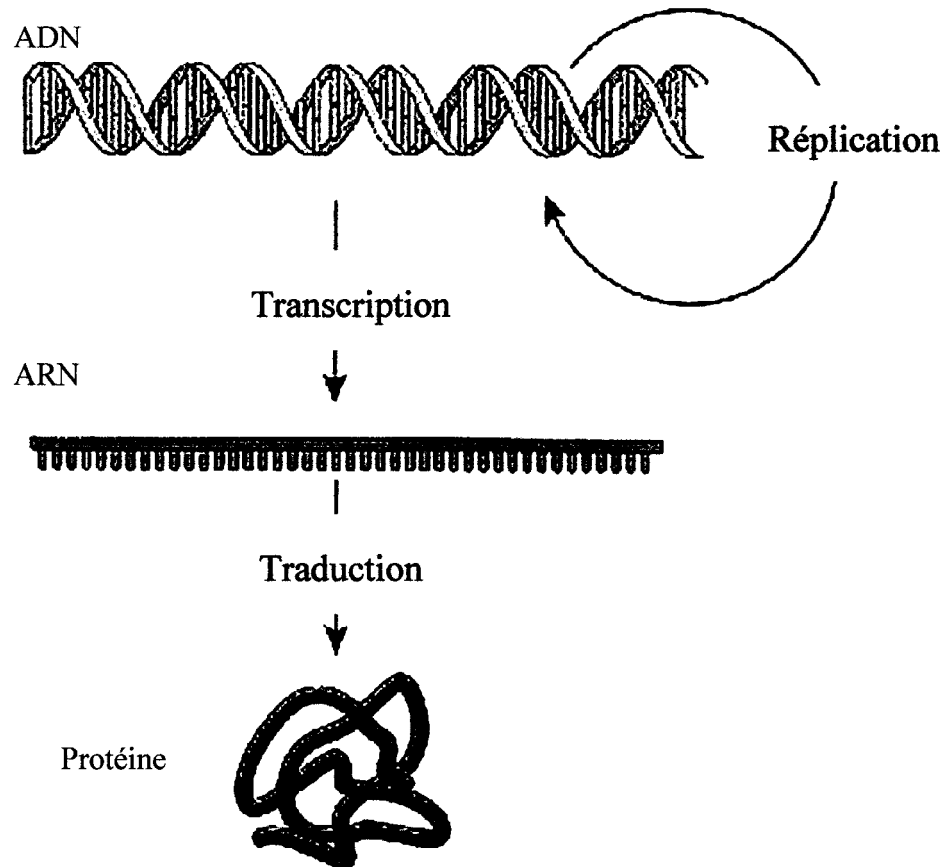
### INTRODUCTION

#### 1.1 Les acides aminés et la nutrition

Le dogme central de la biologie moléculaire actuellement accepté et qui décrit la transmission et l'expression de l'information génétique est montré dans la figure 1.1. Cette figure nous montre l'acide désoxyribonucléique (ADN) dans lequel l'information est stockée sous forme de gène, et est transmise à la descendance par la réplication de ce même ADN. L'expression de cette information se fait, dans un premier temps, par la transcription de l'ADN en acide ribonucléique messager (ARNm). De ces ARNm, les ribosomes traduisent l'information en une séquence d'acides aminés que l'on appelle protéine. Les processus de réplication, de transcription et de traduction sont complexes et sont décrits dans plusieurs ouvrages de qualité tel que *Biochemistry, Second Edition*, de Voet & Voet.

Toute protéine est faite d'un polymère d'acides aminés. Ces acides aminés sont au nombre de 20 et constituent les "blocs Lego" dont la vie se sert pour fabriquer une machinerie cellulaire qui est responsable de toutes les fonctions métaboliques, structurales et régulatrices. Une cellule, qu'elle soit procaryote ou eucaryote, doit maintenir une certaine quantité d'acides aminés libres dans son cytoplasme pour assurer la bonne production de toutes ses protéines. *Escherichia coli*, une entérobactérie de type





**Figure 1.1: Dogme central de la biologie moléculaire.** Cette représentation montre l'ADN dans lequel est regroupé toute l'information génétique sous forme de gène. Cet ADN est répliqué pour donner une copie à chaque cellule fille. Il s'exprime par la transcription en ARN. Les ARN messagers sont ensuite traduits en protéine.

Gram négatif, possède la machinerie métabolique pour synthétiser tous les acides aminés à partir d'un milieu minimal M9 qui contient essentiellement une source de carbone (glucose ou glycérol) et une source d'azote ( $\text{NH}_4\text{Cl}$ ). D'autres organismes, comme l'être humain, ne peuvent synthétiser que 10 acides aminés sur 20. Les acides aminés ne pouvant pas être synthétisés (Arg, His, Ile, Leu, Lys, Met, Phe, Thr, Trp, Val) sont dits essentiels et doivent par conséquent être fournis par la diète (Voet & Voet 1995).

Les animaux d'élevage ont pratiquement les mêmes besoins en acides aminés essentiels (AAE) que nous. Cependant, les besoins en chacun des AAE ne sont pas tous équivalents. Par exemple, la croissance du porc est limitée principalement par l'apport en lysine et en thréonine, tandis que celle de la vache laitière est limitée par l'apport en méthionine et en lysine et d'une façon moindre par la thréonine et la leucine (Santos et al. 1998; Pettigrew et al. 1997). La production de lait de ces vaches est également limitée par sensiblement les mêmes AAE.

La vache laitière est un animal polygastrique, donc possédant un estomac à plusieurs compartiments. Le premier compartiment de l'estomac s'appelle le rumen et est peuplé d'une importante population bactérienne. Ces microorganismes servent à la dégradation de la plus grande partie de la diète ingérée (cellulose, protéine, acide phytique). Cette flore microbienne métabolise 70% des protéines ingérées par la bête. En retour, cette flore produit une quantité comparable de protéines sécrétées qui sont métabolisées par la vache. Des études conduites au cours des années 60 ont démontré que les protéines produites dans le rumen étaient suffisantes pour la croissance et la

production d'une vache laitière produisant 4500 kg de lait par année (Virtanen 1966). Cependant, le rendement d'une vache moyenne au États-Unis a plus que doublé durant les 40 dernières années. Maintenant, le défi est d'assurer un apport protéique suffisant pour des vaches produisant de 9000 à 14000 kg de lait annuellement. Le contenu en protéines produites par les microorganismes du rumen, en plus d'être insuffisant, contient un profil pauvre en AAE. Pour assurer un tel apport protéique à l'animal, on doit soit : 1- augmenter la quantité de protéines riches en AAE de la diète qui échappe à la protéolyse et la métabolisation dans le rumen ou soit : 2- augmenter la qualité des protéines produites par la microflore du rumen (Santos et al. 1998).

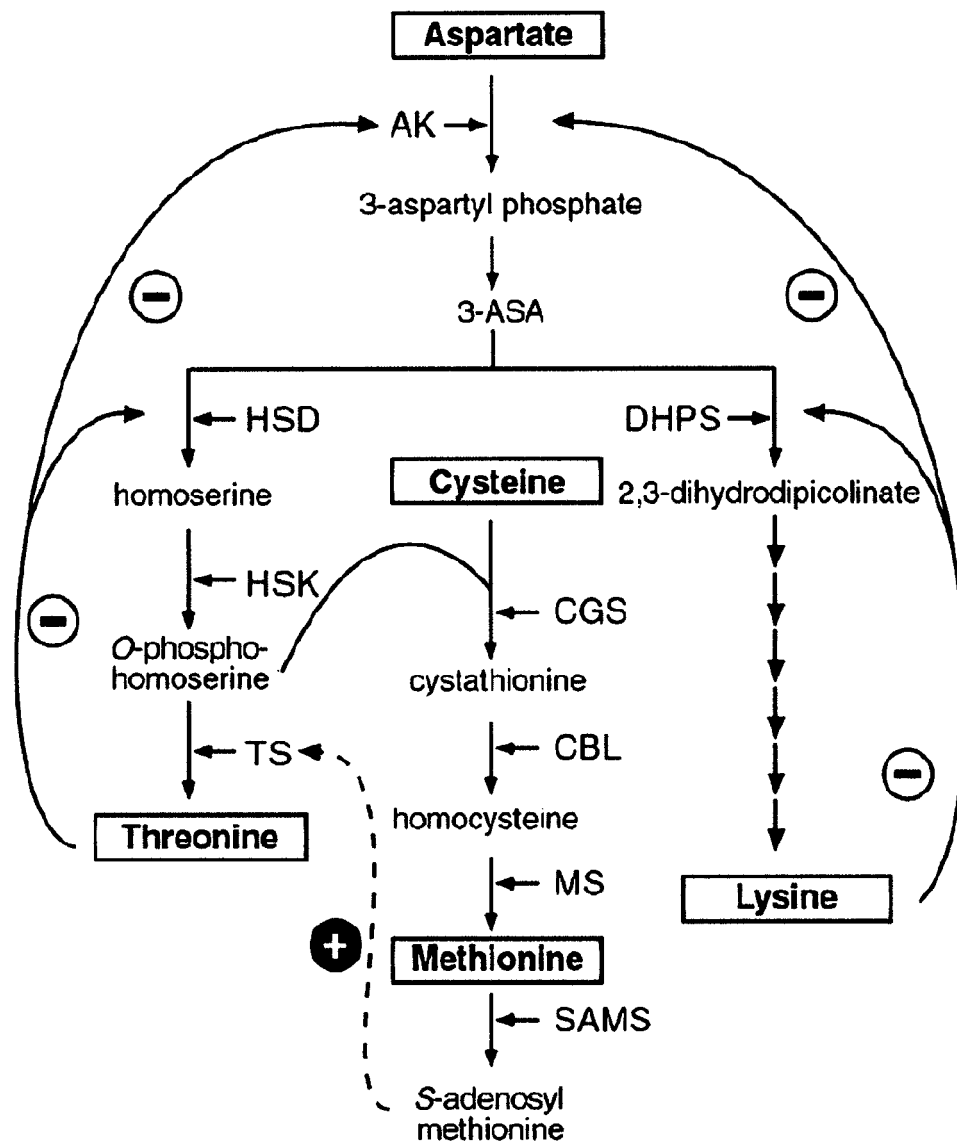
## **1.2 Les solutions pour une meilleure alimentation des animaux d'élevage**

Durant la dernière décennie, les solutions proposées par l'industrie agroalimentaire pour la nutrition bovine ont impliqué des acides aminés libres encapsulés. Les capsules sont faites pour résister à l'environnement du rumen et pour être dégradées dans l'intestin de la vache, délivrant ainsi les acides aminés directement pour l'absorption. Des compagnies multinationales telles que Aventis, Novus International, Mitsui & Co., Nippon Soda Company, Rhone Poulenc et Degussa-Huls occupent près de 90% de ce marché global de 2.25 Milliards de US\$, ce qui laisse très peu de place pour d'autres joueurs. Le marché de la méthionine de synthèse représente un marché de près de 600 millions de dollars américains par année pour un des plus gros joueurs, Aventis. (<http://www.an.aventis.com/>)

La biotechnologie offre pour les compagnies canadiennes une fenêtre d'opportunité unique pour s'insérer dans le marché des AAE pour les animaux d'élevage. Les produits proposés, pour être concurrentiels avec les acides aminés encapsulés, doivent être bon marché et offrir une efficacité nutritive comparable. Trois solutions sont en ce moment explorées : 1- L'ingénierie métabolique; 2- L'expression hétérologue d'une protéine naturelle riche en AAE; 3- L'expression hétérologue d'une protéine de design *de novo* riche en AAE.

### 1.2.1 L'ingénierie métabolique

La première option, l'ingénierie métabolique, consiste à utiliser les outils de la biologie cellulaire pour modifier les flux métaboliques afin de favoriser l'accumulation cytoplasmique de certains acides aminés d'intérêt. Cette approche peut être utilisée par exemple sur une plante fourragère où la proportion de AAE est faible, pour en augmenter le contenu en méthionine ou en lysine et en augmenter d'autant la valeur nutritive. On vise principalement l'augmentation de ces deux acides aminés, car ils sont les plus limitants pour la croissance et la lactogénèse et donc d'un grand intérêt économique. Cette approche est également organe-spécifique, c'est-à-dire que l'expression ou l'inhibition d'un enzyme du métabolisme peut se faire dans un organe de la plante en particulier, par exemple, la graine pour les céréales, l'organe le plus intensivement récolté, ou la feuille dans le cas des plantes fourragères (comme la luzerne ou le trèfle).



**Figure 1.2: La voie générale de biosynthèse de la famille de l'aspartate, comprenant les voies métaboliques de synthèse des AAE méthionine, lysine et thréonine.** Les traits pleins avec le signe (-) représentent des rétro-inhibitions par le produit final de la voie biosynthétique. Les traits pointillés avec le signe (+) représentent une activation de l'enzyme. Tiré de Galili et Höfgen 2002.

Cette stratégie a été utilisée avec succès par plusieurs équipes de recherche (Mazur et al. 1999; Falco et al. 1995; Heremans and Jacobs 1995). L'approche de Falco, par exemple, utilise une dihydropicolinate synthase (DHPS) bactérienne qui est insensible à la rétro-inhibition par la lysine (figure 1.2). La DHPS native s'inactive lorsqu'un certain taux de lysine est atteint dans la cellule. L'expression de l'enzyme bactérienne dans cette cellule permet l'activation constitutive de la voie de synthèse de la lysine et permet par conséquent une plus grande accumulation de cette AAE. Cette technique a permis la production d'un soya ayant un taux de lysine deux fois supérieur à la normale. Des approches semblables pour la méthionine ont également été rapportées (Kim et Leustek 2000; Boerjan et al. 1994). Cependant, l'augmentation du taux de méthionine a été accompagnée de changements morphologiques et de sécrétion de méthanthiol volatil qui donne à la plante une odeur d'ail. Plusieurs autres projets théoriquement valables ont mené à aucune augmentation significative de la méthionine libre, suggérant la présence de facteurs de régulation non identifiés existant dans certaines espèces végétales (Galili 1995; Karchi et al. 1993).

Globalement, les augmentations les plus significatives en lysine ou en méthionine ont toujours été accompagnées par des troubles de germination et de croissance. Il semble que les végétaux utilisent les acides aminés libres comme des messagers cellulaires. Ces phénomènes ne sont que partiellement compris pour l'instant. Force est de constater que la plupart des tentatives de modification de l'équilibre de ces acides aminés provoquent une réaction imprévisible de la plante. De plus, le métabolisme du soufre, donc celui de la méthionine, est très dépendant de la disponibilité, de l'absorption et du transport du

soufre à partir du sol. Des limitations de l'accumulation de méthionine en lien avec la disponibilité en soufre total ont déjà été relevées dans la littérature (Tabe et al. 2002). Jusqu'à présent, l'ingénierie métabolique a permis d'amasser une expertise assez importante pour commencer à envisager la mise au point de produits applicables en alimentation animale, comme le démontre l'abondance de brevets dans le domaine.

### 1.2.2 L'expression hétérologue d'une protéine naturelle riche en AAE

La deuxième approche comprend l'expression par un organisme d'intérêt d'une protéine, provenant d'un autre organisme, naturellement riche en méthionine ou en lysine. Ces protéines servent de puits ("Sink Protein") pour les AAE afin qu'ils échappent au métabolisme. En retirant les AAE du métabolisme, on les met en stockage et on permet au métabolisme de les synthétiser sans provoquer une rétroinhibition. Encore une fois, des facteurs économiques font que seules la Met et la Lys sont exploitées dans ce domaine. Ces protéines doivent être accumulées jusqu'à la consommation de l'organisme par l'animal. Pour ce type d'application, les organismes utilisés pour porter le transgène peuvent être de deux types : 1- Bactéries du rumen; 2- Plantes fourragères ou céréales (de Lumen et al. 1999; Hefford 1997).

Les bactéries naturellement retrouvées dans le rumen peuvent être modifiées pour produire une protéine enrichie en méthionine ou en lysine dans leur cytoplasme. Ces bactéries sont ensuite inoculées et se reproduisent dans le rumen de la vache. Une partie de ces bactéries quitte le rumen avec les aliments ingérés. Elles sont ensuite digérées, ce

qui rend disponible à l'assimilation les AAE de la protéine enrichie (Teather 1985). Les principales espèces ciblées par ces techniques sont *Prevotella ruminicola*, *Butirivibrio fibrisolvens* et *Streptococcus bovis*. Ces dernières sont les plus importantes dans la microflore ruminale (Forano et al. 2000).

Les plantes fourragères et les céréales sont utilisées de la même façon que les bactéries du rumen : on y exprime la protéine enrichie dans l'organe consommé par la vache. Ainsi, la diète de l'animal contient un taux d'AAE plus élevé, ce qui permet de réduire la quantité d'acides aminés synthétiques à ajouter et donc de réduire les coûts de production. Les principaux végétaux utilisés sont le colza, le soja, le maïs et la luzerne.

Très peu de protéines naturelles ont un taux significativement plus élevé en lysine ou en méthionine. La plupart des démonstrations de plantes transgéniques, exprimant une protéine riche en méthionine, ont été réalisées avec entre autre l'introduction du gène de l'albumine 2S de la noix du Brésil, de celui des zéines de maïs et de celui de l'albumine de tournesol (Tabe et al. 1998; Chui et al. 1995; Kortt et al. 1991; Altenbach et al. 1987). Les résultats de ces démonstrations ont été variables. Dans certains cas, la quantité de méthionine totale était augmentée de 5 à 7 fois. Dans d'autres cas, l'expression de la protéine riche en méthionine était corrélée avec la diminution des protéines endogènes riches en méthionine, ce qui semble indiquer une limite de disponibilité du soufre (Chakraborty et al. 2000; Muntz et al. 1998). Certains projets ont également fait face à des problèmes d'immunogénicité de la protéine exprimée, spécialement avec l'albumine 2S de la noix du Brésil (Alcocer et al. 2002). Les seules démonstrations convaincantes



avec des protéines riches en lysine ont été faites avec un homologue modifié de l'inhibiteur de la chymotrypsine-2 et avec la protéine S-VSP (Guenoune et al. 2002; Roesler et al. 2001).

Des approches binaires, combinant l'ingénierie métabolique et l'expression hétérologue, ont connu un succès très intéressant. Jung et Falco ont fait exprimer dans le maïs la protéine riche en lysine BHL8 avec un rendement de 3-6% des protéines totales de la graine, qui, lorsque combiné avec l'expression d'une DHPS bactérienne (voir la section 1.2.1 et la figure 1.2), a fait augmenter la quantité de lysine totale à 0.7% du poids sec de la graine, comparé à 0.2% pour le maïs ordinaire (Jung et Falco 2000). Plusieurs multinationales ont breveté des systèmes binaires pour l'application dans des fourrages transgéniques. Ces procédés sont en voie de commercialisation (Beach et al. 2000; Jung et al. 1998; Falco et al. 1996).

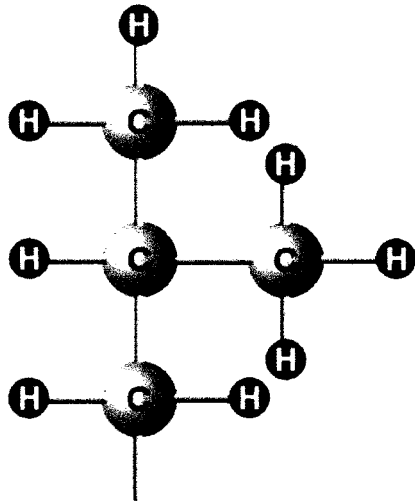
### 1.2.3 L'expression hétérologue d'une protéine de design *de novo* riche en AAE.

La troisième approche implique l'expression par un organisme, comme une plante fourragère ou une bactérie du rumen, d'une protéine nouvelle (qui ne provient pas de la nature) et qui possède un biais important et **contrôlé** en AAE. Cette protéine, comme la précédente, sert de puit pour les AAE. Contrairement aux deux précédentes approches qui visaient très souvent un seul acide aminé à la fois, cette approche possède l'avantage d'être malléable, car le contenu de la protéine peut être modifié et adapté au profil en AAE nécessaire à l'espèce et/ou à la diète visée.

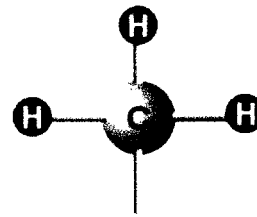
Les premiers efforts d'expression de protéines synthétiques avec un haut taux d'AAE *in vivo* dans des bactéries n'ont pas donné de hauts rendements en protéine, vraisemblablement dû à un manque de stabilité structurale du peptide final (Beauregard et al. 1994; Jaynes et al. 1985; Doel et al. 1985).

Le design de protéines en général a été caractérisé jusqu'à maintenant par la simplicité des constructions et par l'absence quasi-absolue de méthionine (Offredi et al. 2003; Walsh et al. 2001; Johansson et al. 1998; Blake et al. 1998; Schafmeister et al. 1997; Jiang et al. 1997; Brive et al. 1997). Les designers de protéines évitent la méthionine en raison de sa pauvre capacité à spécifier un repliement et de sa faible énergie libre de transfert au solvant. En effet, la chaîne latérale de la méthionine possède deux liens flexibles de plus que la leucine (un des acides aminés hydrophobes les plus communs) comme on peut le voir à la figure 1.3. On estime que cette grande flexibilité entraîne, lors de la fixation d'un rotamère unique dans la structure de la protéine native, un coût additionnel en énergie libre entropique de  $0.8 \text{ kcal mol}^{-1}$ . De plus, dû à une plus grande polarité que la plupart des autres acides aminés hydrophobes, le transfert de la chaîne latérale de la méthionine du solvant vers le cœur hydrophobe lors du repliement amène un gain en énergie libre de transfert au solvant. Par exemple, ce gain équivaut approximativement à  $0.6 \text{ kcal} \cdot \text{mol}^{-1}$  de moins que pour l'enfouissement d'une leucine (Gassner et al. 2003).

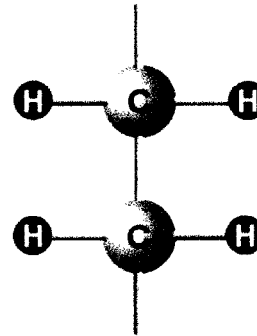
Quelques groupes se sont concentrés sur l'identification des principes de replie-



Leucine (Leu)



S



Méthionine (Met)

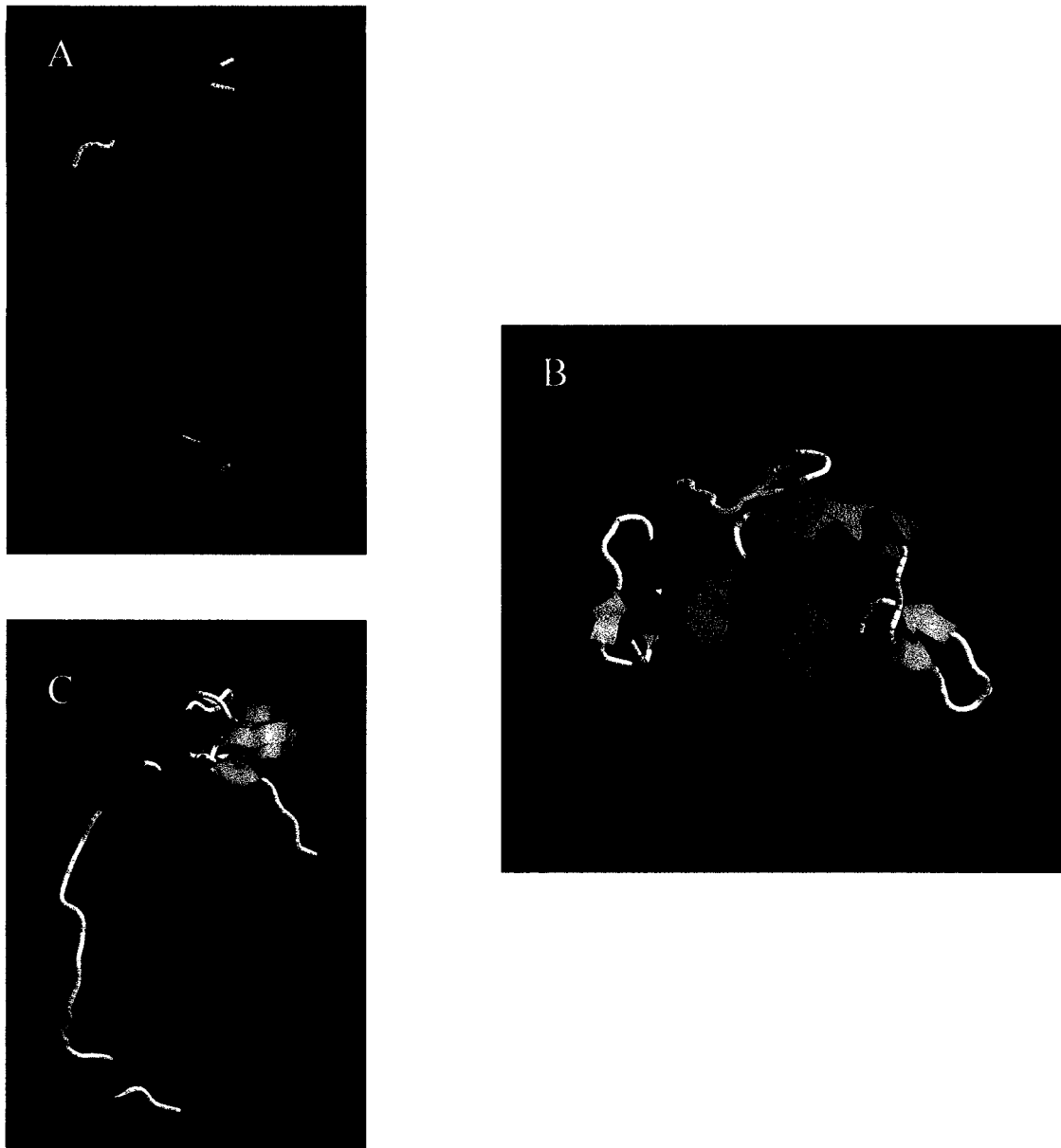
**Figure 1.3: Structure de la chaîne latérale de la leucine et de la méthionine.** Puisque la chaîne latérale de la leucine est ramifiée, elle possède beaucoup moins de liberté que celle de la méthionine.

-ment permettant d'intégrer le plus de méthionines possibles dans de nouvelles séquences d'acides aminés.

Jusqu'à maintenant, aucune expression significative d'une protéine de design *de novo* dans un organisme d'intérêt nutritionnel n'a été réalisée. Le design de protéines nutritives, spécialement celles riches en méthionine, demeure un défi extrêmement actuel et une fenêtre d'opportunité grande ouverte pour les compagnies de biotechnologie canadiennes.

### 1.3 Le design de protéines

Le design de protéines peut se faire de plusieurs façons. Généralement, les designers visent d'abord un repliement spécifique sur lequel ils possèdent des informations venant d'exemples de protéines naturelles qui ont la même topologie. Les lignes directrices au design du motif sont en général minimalistes. Elles se font sur de petites protéines ayant un nombre limité d'interactions non-covalentes, des séquences hautement répétitives, des structures secondaires uniques et une haute symétrie. Des prototypes intéressants ont été caractérisés avec des motifs hélicaux simples tels que les "torsades d'hélices" (Harbury et al. 1998; Ogihara et al. 1997), "les carrées d'hélices" (Walsh et al. 2001; Johansson et al. 1998; Blake et al. 1998; Schafmeister et al. 1997; Jiang et al. 1997; Brive et al. 1997; Kamtekar et al. 1993), "les doigts de zinc" (Pomerantz et al. 1998; Dahiyat et al. 1997; Struthers et al. 1996) (voir la figure 1.4) et quelques autres protéines simples (Malakauskas et al. 1998; Riddle et al. 1997; Lazar et al. 1997;



**Figure 1.4: Les motifs tridimensionnels les plus courants pour des protéines de design *de novo*.** A : Torsade d'hélices. Domaine d'oligomérisation de l'antigène  $\delta$  du virus de l'hépatite [pdb# 1A92]; B : Doigt de zinc. Zif268 protéine de réponse à l'hormone de croissance liée à l'ADN [pdb# 1JK1] ; C : Carrée d'hélices. Rubrerythrine [pdb# 1LKP].

Gu et al. 1995).

On peut voir également dans la littérature quelques rares exemples de protéines riches en structures  $\beta$  (Yang et al. 2003; Xu et al. 2001; Houbrechts et al. 1995; Beauregard et al. 1991). Ces protéines sont pour la plupart peu solubles. Cependant, la bio-informatique, via le design *in silico*, a récemment permis d'obtenir des exemples de réussites (Offredi et al. 2003; Wang et al. 2002; Richardson et al. 2002; de Alba et al. 1999).

Malgré les efforts intenses des équipes de recherche, l'approche minimaliste de la plupart des protéines de design *de novo* mène à un comportement peu semblable à celui d'une protéine naturelle dite "native".

Les protéines naturelles dans leur état natif se replient en une structure unique qui est stabilisée par des interactions hydrophobes, de l'empaquetage spécifique de chaînes latérales (interactions de Van der Waals), des réseaux de ponts hydrogènes et des interactions électrostatiques. On décrit les protéines naturelles natives par les caractéristiques suivantes (Kohn et al. 1998):

- Un déplacement chimique avec une bonne dispersion en RMN, indiquant un environnement chimique unique bien défini pour chaque atome,
- Un état oligomérique défini (monomère, dimère, trimère, etc.),

- Un cœur hydrophobe avec un empaiquetage spécifique, sans vide ni crevasse, indiquer par : un haut facteur de protection à l'échange des hydrogènes amides de la chaîne principale en RMN, une bonne résistance aux protéases et une faible affinité à l'ANSA,
- Une dénaturation thermique et chimique coopérative (sigmoïdale),
- Un grand  $\Delta C_p$  de dénaturation, indicatif du changement de l'exposition de surface hydrophobe au solvant lors de la dénaturation.

Contrairement aux protéines naturelles, on retrouve parmi les protéines de design *de novo* de nombreux cas d'état "globule fondu" (molten globule). Cet état a été intensivement étudié et il a des caractéristiques biophysiques à peu près contraires à celles énumérées ci-haut.

Le globule fondu a été décrit comme étant un état partiellement plié ou intermédiaire, dans lequel l'essentiel des structures secondaires est formé, mais où l'empaquetage fin des chaînes latérales n'est pas optimisé. Le cœur hydrophobe est fluide et donne plusieurs populations de conformères d'énergie équivalente en échange rapide. Ces protéines ne se comportent pas bien *in vivo* (problèmes d'expression, d'association non-spécifique et de dégradation protéolytique). L'apomyoglobine, l'alpha-lactalbumine et l'apolipoprotéine E4 ont été les principaux modèles de globule fondu étudiés (Nishimura et al. 2002; Morrow et al. 2002; Permyakov et al. 2000).

Pour remédier à ce problème, les designers de protéines ont développé l'approche par "design négatif". Cette dernière sert à spécifier une conformation unique à la protéine en déstabilisant tous les états qui sont isoénergétiques et accessibles, pour n'en laisser qu'une seule thermodynamiquement probable. Cette approche peut être accomplie par l'incorporation d'interactions entre des chaînes latérales polaires enfouies dans le cœur hydrophobe ou simplement par le résultat d'une complémentarité stérique spécifique des chaînes latérales hydrophobes du cœur. La nécessité d'avoir une structure atomique tridimensionnelle précise est essentielle. Pour le moment, l'impossibilité d'avoir accès à une telle structure pour MB-1 nous oblige à utiliser une stratégie d'essais et erreurs basée sur le modèle de structure en carré d'hélices attendu pour MB-1 (figure 1.5) et sur les prédictions bio-informatiques de positions des hélices dans la séquence.

Récemment, l'optimisation de ces protéines est maintenant devenue possible par des moyens beaucoup moins fastidieux. Deux approches ont fait leurs preuves : les algorithmes de prédiction de repliement et l'évolution dirigée.

Les algorithmes de prédiction de repliement ont émergé au début des années 90. Cependant, ils ont été vraiment développés en 1997, principalement par l'équipe de Stephen Mayo. Il a, pour la première fois, inclus dans ses algorithmes toutes les parties de la protéine : le cœur hydrophobe, les surfaces exposées au solvant et les régions à l'interface (Dahiyat et al. 1997a; Dahiyat et al. 1997b).

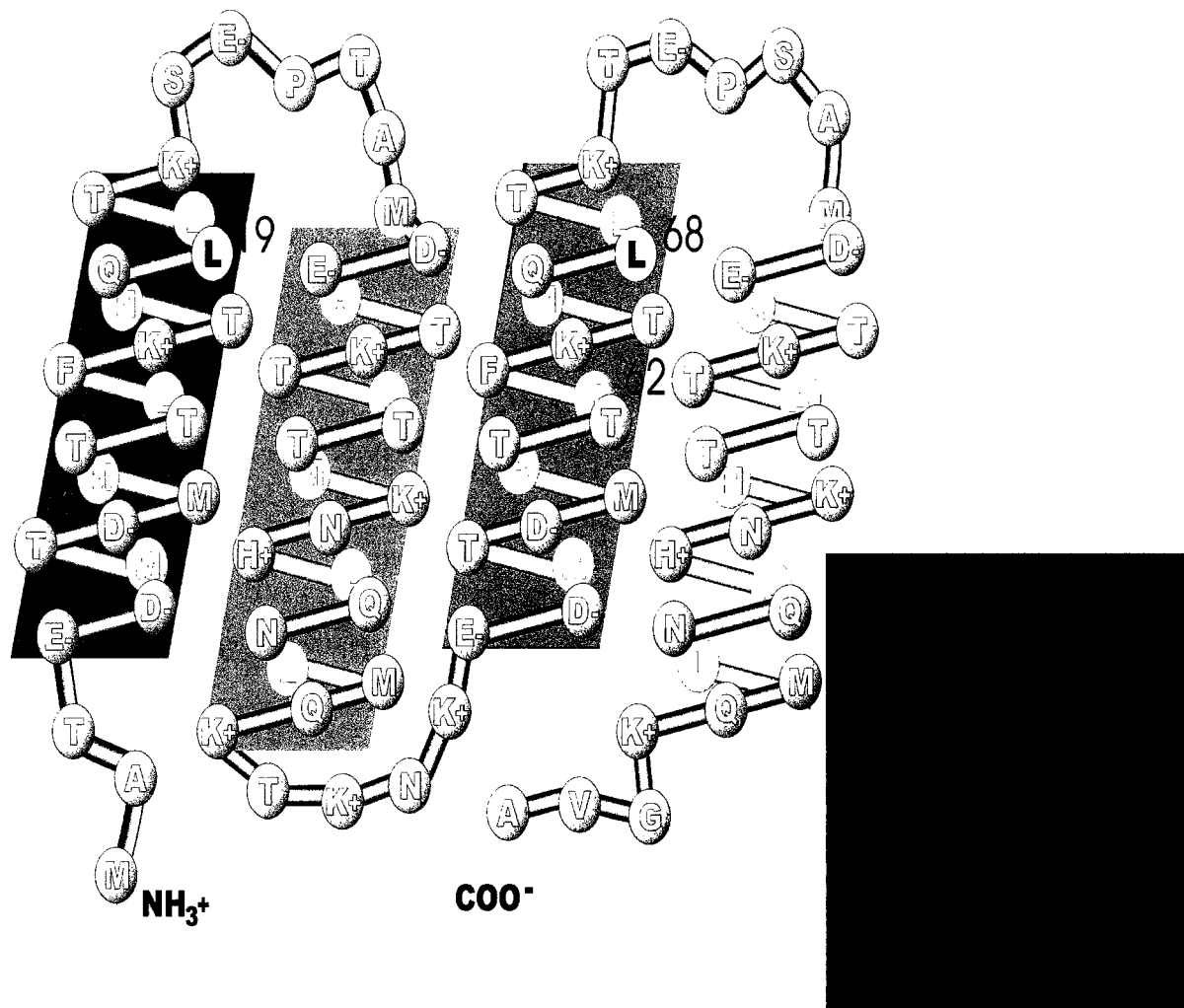


L'évolution dirigée a également débuté durant les années 90. Elle implique une accélération du processus de sélection naturelle basé sur l'apparition aléatoire d'erreurs parfois bénéfiques. Elle peut se faire de façon semi-rationnelle (en ciblant des positions ou des régions précises) ou de manière complètement aléatoire (sur l'ensemble du gène). Des exemples récents dans la littérature illustrent bien chacune des méthodes (Chu et al. 2003; Wang et al. 2000).

#### **1.4 Le projet " Milk Bundle-1 "**

En appliquant le design de protéine à la nutrition animale, notre équipe de recherche a innové en 1995 avec la création de la protéine "Milk Bundle-1" (MB-1). Cette protéine contient 56% d'AAE et est conçue pour se replier en carrée d'hélices (figure 1.5). On peut voir également à la figure 1.5 sa séquence en acides aminés dont les proportions en AAE correspondent aux besoins des vaches laitières de haute performance, soit 14% méthionine, 12% lysine, 21% thréonine et 10% leucine. Il a été démontré que cette protéine a un cœur hydrophobe bien replié ainsi qu'une faible affinité pour une sonde hydrophobe, le 8-anilino-1-naphthalenesulfonic acid (ANSA) (Beauregard et al. 1995). Son comportement et son niveau d'expression *in vivo* étaient de loin supérieurs à ceux d'autres protéines semblables (MacCallum et al. 1997). Il a été démontré que MB-1 est une protéine-puits pour les AAE. De plus, elle se différencie de ses compétiteurs naturels par le contrôle parfait des proportions de chaque AAE.

La caractérisation initiale de MB-1 a montré cependant une faible température de dénaturation (39°C) ainsi qu'une faible résistance aux protéases (MacCallum et al. 1997).



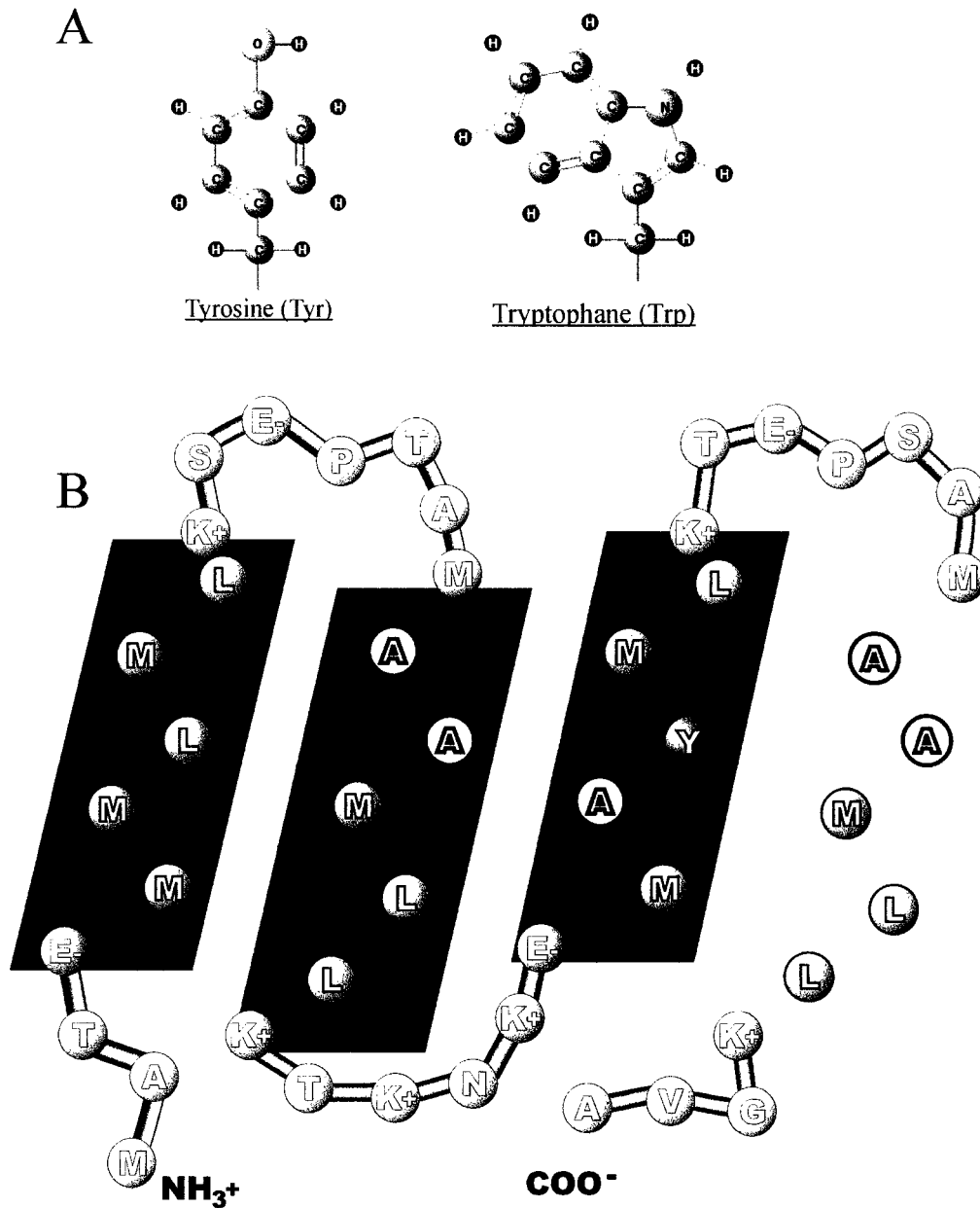
**Figure 1.5: Représentation de MB-1 avec la position prédite de ses hélices.** Sa face hydrophobe se trouve à l'arrière-plan tandis que sa face hydrophile au premier plan. Sont indiquées les positions Leu19, Tyr62 et Leu68. En encadré, le modèle à ruban d'un carré d'hélices, structure attendue pour MB-1.

Ces résultats suggèrent que MB-1 ne s'accumulerait pas dans le cytosol d'une bactérie du rumen ou d'une cellule de plante transgénique, car elle serait sensible à leurs protéases (Huang et al. 1994; Liao 1993; Parcell et al. 1989).

### **1.5 Exemple d'améliorations de MB-1 : MB-1Trp**

Étant donné la faible thermorésistance de MB-1, la fabrication de variants plus performants, soit avec un  $T_m$  et un temps de demi-vie face aux protéases plus grands, a été entreprise. Notre hypothèse de départ était que l'amélioration de l'effet hydrophobe devrait nous permettre d'obtenir une protéine ayant une plus grande stabilité thermique et par conséquent, une meilleure résistance aux protéases. L'effet hydrophobe est maximisé lorsque les chaînes latérales regroupées dans le cœur hydrophobe ont un moment dipolaire très faible (une hydrophilicité faible). MB-1 possède une tyrosine en position 62 servant comme sonde spectroscopique. Son groupement hydroxyle polaire réduit la force de l'effet hydrophobe local. De plus, il n'a aucun partenaire à proximité pour faire des ponts hydrogènes comme illustré à la figure 1.6 (5 alanines l'entourent et font une poche pour accueillir la grosse chaîne latérale), ce qui laisse un groupement polaire insatisfait et non compensé par une force d'attraction (pont hydrogène, pont salin) à l'intérieur du cœur. Le mutant MB-1Trp a été obtenu en substituant la tyrosine en position 62 pour un tryptophane, dont les chaînes latérales respectives peuvent être visualisées à la figure 1.6A. Ce changement a amené une nette amélioration des caractéristiques de la protéine, soit l'augmentation de la température de dénaturation de 39 à 55°C et l'augmentation de la résistance aux protéases jusqu'à un niveau comparable à celui d'une protéine naturelle

globulaire de même taille (ex : cytochrome C) (Gagnon et al. 2000). Cette amélioration a été attribuée à l'optimisation de l'effet hydrophobe de MB-1 par le retrait du groupement hydroxyle et par l'augmentation du volume du cœur engendré par la différence de taille de la chaîne latérale de la Tyr ( $187 \text{ \AA}^3$ ) par celle du Trp ( $217 \text{ \AA}^3$ ) (voir la figure 1.6).



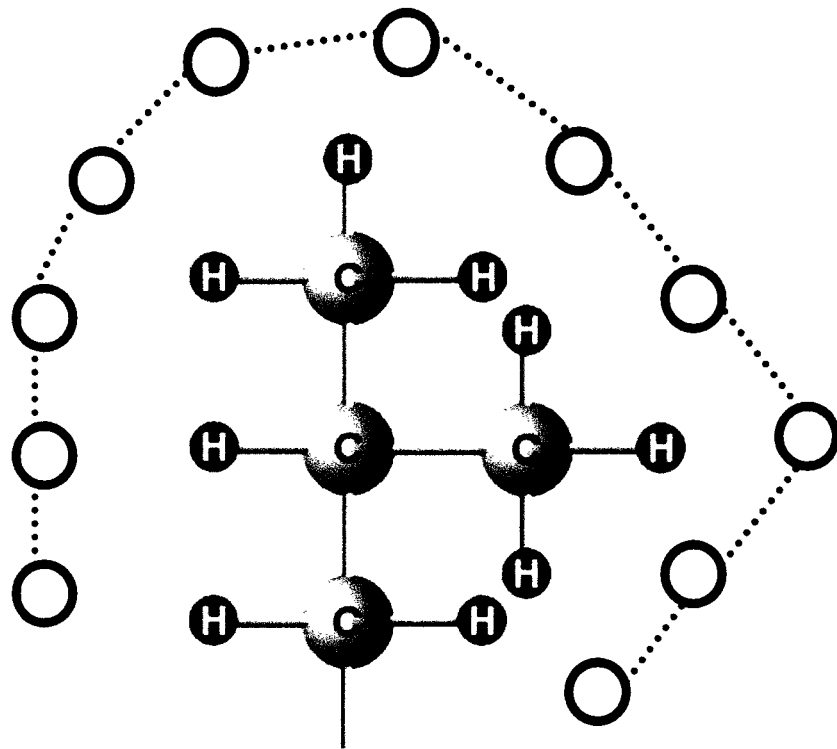
**Figure 1.6: A) Chaînes latérales de la tyrosine et du tryptophane. B) Représentation de MB-1 avec seulement les boucles et les acides aminés participant au cœur hydrophobe. En rouge, le seul résidu tyrosine en position 62 entouré de 5 alanines qui lui font une niche d'accueil dans le cœur.**

## 1.6 La stratégie de cette étude

La principale force qui pousse les protéines à se replier en milieu aqueux est l'effet hydrophobe. Il contribue de plusieurs centaines de  $\text{kJ mol}^{-1}$  à la stabilité du repliement, comparé aux interactions électrostatiques qui ne contribuent que de quelques  $\text{kJ mol}^{-1}$ . Donc, l'optimisation de cet effet hydrophobe est critique pour la stabilité d'une protéine (Wu et al., 1997; Cordes et al., 1996; Richards et al., 1993; Ponder et al., 1987).

L'effet hydrophobe s'observe par le regroupement des chaînes latérales hydrophobes les unes contre les autres lors du repliement d'une protéine, pour les isoler du solvant. Ce phénomène rappelle le regroupement des molécules d'huile pour former des gouttelettes dans l'eau. Les chaînes latérales hydrophobes que la protéine n'arrive pas à enfouir dans le cœur se retrouvent donc exposées au solvant. Les molécules de solvant doivent alors s'organiser en cage (clathrate) autour de ces chaînes non-polaires, comme l'illustre la figure 1.7. L'immobilisation de molécules de solvant a un coût entropique déstabilisant pour la protéine.

Dans le but d'optimiser la contribution stabilisante de l'effet hydrophobe dans la protéine MB-1Trp, les leucines en position 19 et 68, qui sont des résidus hydrophobes exposés au solvant (figure 1.5), seront toutes deux substituées par des résidus polaires qui n'entraînent pas l'immobilisation d'une clathrate d'eau. La contribution stabilisante du retrait d'un groupement méthylène de la surface exposée au solvant d'une protéine a été évaluée à  $-3 \text{ kJ mol}^{-1}$  (Voet & Voet 1995).



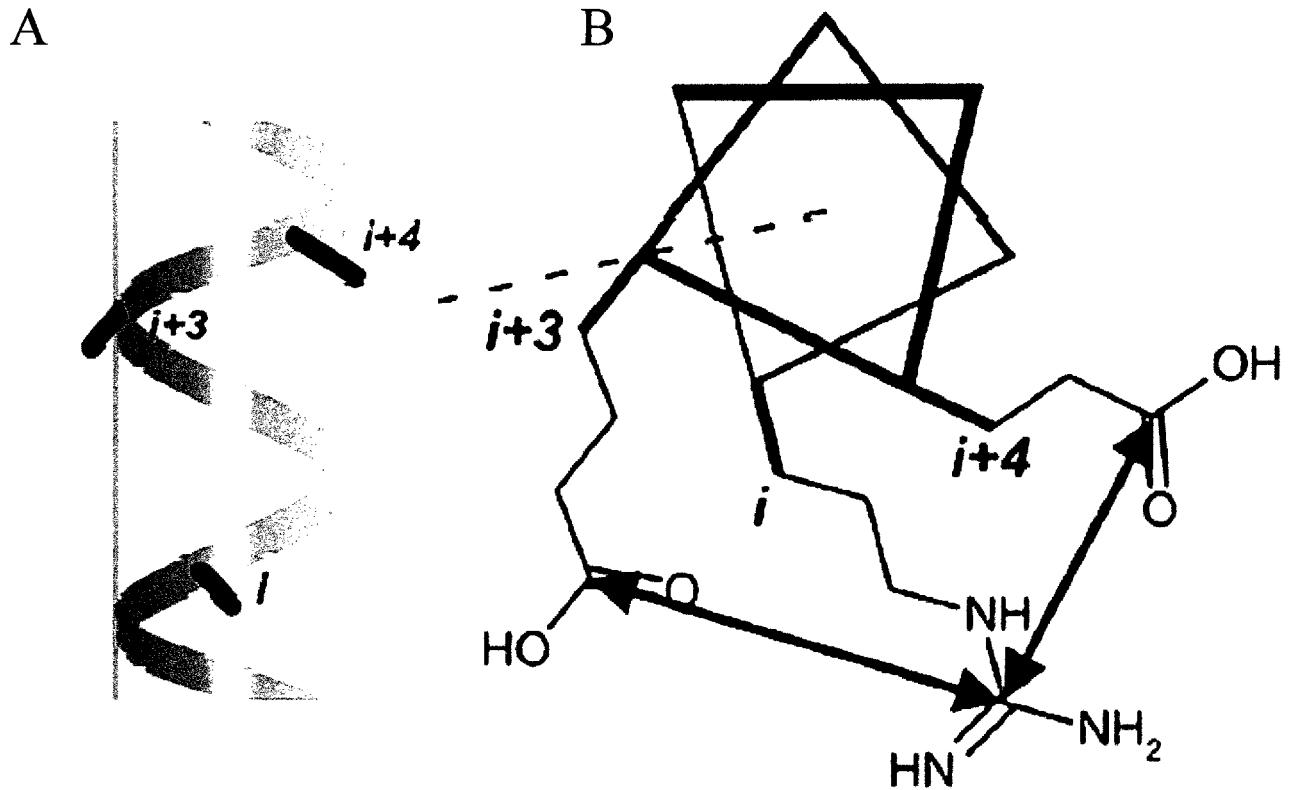
## Leucine (Leu)

**Figure 1.7: Clathrate de molécules d'eau, représentées ici par les sphères blanches, organisées autour d'une chaîne latérale d'une leucine. Cette organisation et l'immobilisation des molécules d'eau entraîne un coût entropique qui déstabilise la protéine dans son état natif.**

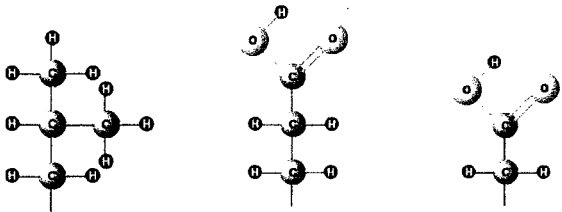
Le choix d'un acide aminé chargé négativement comme substituant a été motivé par la possibilité de former des ponts salins intra-hélicaux. Il a été montré que les ponts salins peuvent avoir un impact stabilisant important sur une protéine (Kumar et al. 2002; Takano et al. 2000; Loladze et al. 1999; Merz et al. 1999). On sait qu'une chaîne latérale en position  $i$  peut former un pont salin à l'intérieur d'une hélice avec les chaînes latérales en positions  $i+3$ ,  $i+4$ ,  $i-3$  et  $i-4$  tel que montré en figure 1.8 (Burkhard et al. 2002). Comme on peut le voir en figure 1.5, la position 19 est entourée par les lysines 15 et 22, des résidus polaires chargés positivement, en  $i+3$  et  $i-4$ , c'est-à-dire en position pour former des ponts salins. La position 68 a des résidus avoisinants similaires et peut, elle aussi, former des ponts salins.

Il a été envisagé d'utiliser l'Asp, qui est un acide aminé polaire chargé négativement, pour remplacer la leucine19 et la leucine68. Bien qu'il ait, tout comme le Glu, une bonne hydrophilicité, il a été écarté en raison de sa capacité à servir de casseur d'hélice dans certains cas (Burkhard et al. 2002), de son manque de propension à former des hélices et de sa différence de volume avec la leucine (Figure 1.9) (Koehl et al. 1999). Certains auteurs ont postulé que les positions exposées au solvant ont généralement une plus grande capacité à accueillir des acides aminés de tailles variées que les positions de cœur. Par conséquent, la différence de volume ne constituerait pas un argument convenable pour justifier le choix de notre substituant. Cependant, dans l'optique de faciliter l'analyse de l'impact des mutations sur la protéine, les auteurs de cette étude ont préféré réduire l'impact de la mutation à un changement de polarité seulement (dans le cas de Leu→Glu) par rapport à un changement simultané de polarité et de taille (dans le





**Figure 1.8: A) Vue latérale d'un polypeptide en conformation hélice- $\alpha$ . B) Vue de haut, le même polypeptide où il est montré les interactions électrostatiques de type ponts salins entre trois chaînes latérales.** Chaque trait de l'heptade représente la chaîne principale d'un acide aminé. En position  $i$  se trouve une chaîne latérale d'arginine et en  $i+3$  et  $i+4$ , des chaînes latérales de glutamate. (Tiré de Burkhard et al. 2002)



The image shows the chemical structures of the side chains for Leucine (Leu), Glutamate (Glu), and Aspartate (Asp). Leu has a branched aliphatic side chain. Glu and Asp have carboxylate side chains, with Glu having one more methylene group than Asp.

Acides aminés	Leu	Glu	Asp
Volume de Van der Waals (Å <sup>3</sup> )*	137	138	106
Propension d'hélice**	-0.38	-0.33	0.27
Indice d'hydrophobicité***	-1.8	2.5	2.5

\*Miller et al., 1987.

\*\*Koehl and Levitt, 1999.

\*\*\*Nozaki and Tanford, 1971 et Levitt, 1976.

**Figure 1.9: Comparaison des caractéristiques de la leucine, du glutamate et de l'aspartate.** On y constate que le glutamate est un homologue polaire de la leucine beaucoup plus ressemblant que l'aspartate.

cas de Leu→Asp).

Donc, les glutamates ont finalement été choisis comme substituants en raison de leur plus grande aptitude à former des ponts salins à l'intérieur d'une hélice et leur grande propension hélicale (Burkhard et al. 2002). De plus, comme on peut le voir à la figure 1.8, la substitution de la leucine pour un glutamate provoque la disparition de deux groupements méthyl de la surface de la protéine, entraînant, par le fait même, une contribution théorique stabilisante de  $-6 \text{ kJ mol}^{-1}$ .

Plusieurs études ont démontré la possibilité d'augmenter la stabilité thermique et la résistance aux protéases d'une protéine soit par une approche d'ajout de ponts salins ou par l'optimisation de l'effet hydrophobe. Le but de mon projet est de vérifier la possibilité d'appliquer ces stratégies à une protéine nutritive *de novo*, MB-1Trp, en utilisant des mutations de surface. Donc, trois protéines seront caractérisées dans cette étude; Leu19Glu, Leu68Glu et Leu19Glu/Leu68Glu, représentant chacune la protéine MB-1Trp avec la mutation leur donnant leur nom respectif. Ces mutants ont été construits par mutagenèse dirigée sur le gène gabarit MB-1Trp auxquels ont été ajouté par ligation une queue de six histidines en C-terminal pour faciliter la purification. Ils ont été exprimés dans *E.coli* et purifiés par différents types de chromatographies d'affinité. Leurs structures secondaires ont été caractérisées par dichroïsme circulaire. L'intégrité de la structure tertiaire a été démontrée par fluorescence du tryptophane et de l'ANSA. L'amélioration des paramètres de stabilité et résistance de la structure a été évaluée par dénaturation thermique des structures secondaires suivi par dichroïsme circulaire et par

dégradation protéolytique face à l'exposition à un mélange de protéases de *Streptomyces griseus*, Pronase E (Sigma-Aldrich #P5147). Les stratégies d'optimisation de l'effet hydrophobe et de l'ajout d'interactions électrostatiques à la surface de la protéine seront donc employées simultanément sur notre modèle. Nos résultats montrent qu'il est possible d'augmenter la stabilité d'une protéine *de novo* malgré les difficultés inhérentes à ce genre d'entreprise, dans ce cas, l'absence d'une structure atomique tridimensionnelle.

## **CHAPITRE II**

### **RÉSULTATS**

#### **2.1 Article scientifique**

Single-point mutations at the surface of MB-1 Trp lead to important changes  
in conformational properties

Maxime Sasseville, Sébastien Claveau, Mylène C. Gagnon, Marc Beauregard\*.

Sous presse dans Journal of Peptide Science.

**Résumé:** Le design de protéine est utilisé pour créer de nouvelles protéines ayant des caractéristiques désirables. Dans notre laboratoire, nous visons la synthèse de protéines contenant un haut taux d'acides aminés essentiels ayant une application potentielle en nutrition animale. Une des limitations auxquelles nous devons faire face dans cette entreprise est l'atteinte d'une stabilité malgré un contenu en acides aminés hautement biaisé. Nous rapportons ici la synthèse et la caractérisation de trois variants de la protéine MB-1Trp dans lesquelles deux Leu exposées au solvant ont été remplacées par des Glu permettant la formation de nouveaux ponts salins à la surface de la protéine. Quoique les deux mutations étaient crû similaires (c'est-à-dire, la même mutation dans un environnement local similaire), elles semblent avoir des répercussions différentes sur MB-1Trp tel que montré par dichroïsme circulaire de UV lointain, par dénaturation thermique, par fluorescence et par mesures de la résistance protéolytique. Pour la mutation Leu68Glu, une augmentation de 6°C de la température de dénaturation de la protéine a été observée. Curieusement, la mutation Leu19Glu a mené à une diminution de la température de dénaturation et une modification de la structure tertiaire.

Single-point mutations at the surface of MB-1Trp lead to important changes  
in its conformational properties

Maxime Sasseville, Sébastien Claveau, Mylène C. Gagnon, Marc Beauregard.

Authors' affiliations: Groupe de recherche en énergie et information biomoléculaires.

Département de Chimie-*Biologie*. Université du Québec à Trois-Rivières, Trois-Rivières,  
Québec, Canada.

Correspondence to:

Marc Beauregard

Département de Chimie-Biologie

Université du Québec à Trois-Rivières

Trois-Rivières, Qc G9A 5H7

E-mail: marc\_beauregard@uqtr.ca.

Telephone: (819) 376-5011 #3354. Fax: (819) 376-5084.

**Running title:** Surface mutations on MB-1Trp

**Keywords:** Protein design; hydrophobic effect; surface salt bridge; essential amino acids; nutritional proteins

**Summary:** Protein design is currently used for the creation of new proteins with desirable traits. In our lab we focus on the synthesis of proteins with high essential amino acid content having potential applications in animal nutrition. One of the limitations we face in this endeavour is achieving stable proteins despite a highly biased amino acid content. We report here the synthesis and the characterization of three variants of MB-1Trp in which two solvent-exposed Leu have been replaced by Glu allowing for the formation of new salt bridges at the surface of the protein. Though both mutations were expected to be similar (i.e. same mutation in a comparable local environment), they appear to have different effects on MB-1Trp as shown by far-UV circular dichroism, thermal denaturation, fluorescence and proteolytic resistance measurements. For the mutation Leu68Glu, an increase in the protein melting temperature of 6°C was observed. Surprisingly, the mutation in position Leu19Glu led to a decrease in melting temperature and a modification of tertiary structure.



**ABBREVIATIONS**

ANSA: 8-Anilino-1-naphthalene sulfonic acid; BCA: bicinchoninic acid; CD: circular dichroism; EAA: essential amino acids; Glu: glutamate; Leu: leucine; MB-1: milk-bundle-1; Met: methionine; Lys: Lysine; PAGE: polyacrylamide gel electrophoresis; Trp: tryptophane; Tyr: tyrosine.

## INTRODUCTION

The expense associated with feed additives used for animal production has led to the study of intracellular production of high-quality protein by transgenic crops and other organisms as a means of obtaining efficient and less costly sources of essential amino acids (EAA), mainly Met and Lys (1-4). Two main trends are actually being explored: metabolic engineering and heterologous expression. The former includes the manipulation of free amino acids metabolic fluxes to promote accumulation of EAA (for a review see 5). The latter option can be achieved by either the expression and accumulation of a naturally occurring protein enriched in EAA or by the expression of a *de novo* designed protein with a controlled content in EAA.

Natural proteins rarely come with an strong bias in Met and Lys. However, some success of enriched protein expression in plants have been reported using sunflower seed albumin and engineered barley chymotrypsin inhibitor-2 (6, 7). Recently, soybean and sunflower albumins have been chosen on the base of their high methionine contents for the development of transgenic crops by major agro-biotechnology companies (8, 9).

*De novo* design of protein offers a great advantage for such application: the EAA content in designer proteins can be controlled to fit the needs of any animal production type. One of the most important EAA for animal performance is methionine. Positioning methionine side chains in a protein scaffold is more challenging than for other EAA: protein designers avoid methionine mainly because of its poor fold specifying characteristics and its lower solvent transfer free energy (10). Further, methionine is a target for oxidation, which may destabilize a methionine enriched hydrophobic core (11,

12). Not surprisingly, methionine has been avoided when designing the core of most designer proteins reported to date (13-17).

Moving from repetitive genetic constructions (18) to rational design of a putative alpha-helical bundle, we generated a novel protein, MB-1, that contains a very high content in EAA, amongst which methionine is highest in proportion (19). Its behaviour and expression levels *in vivo* were found to be far superior to those obtained from earlier attempts in this field of research (20). However, investigation of MB-1 properties did reveal some flaws. Its thermal denaturation had a low cooperativity and its melting temperature was only of 39°C. Moreover, its protease resistance at physiological temperatures was limited, possibly due to partial unfolding (20). Such results suggest that MB-1 would not accumulate in transgenic organisms, where it would be exposed to endogenous proteases (21-23).

It is well known that hydrophobic interactions play an important role in protein folding and stability (24-27). The optimization of the global hydrophobicity and packing of non-polar side-chains in the core, while increasing solvent transfer free energy, has proven its value for conformational stabilisation of proteins (28-30). In order to further improve conformation stability, one should maximize the hydrophobic effect not only by working on the core, but also at the surface of a protein. In this respect, the solvent exposure of hydrophobic side chains should be minimised, and vice versa for hydrophilic residues.

Using MB-1 as a template a second generation mutant protein named MB-1Trp was designed using this strategy. A single substitution of a Tyr for a Trp in the hydrophobic core, which increases the hydrophobicity and bulk in the core, lead to an

impressive improvement of melting temperature, from 39°C to 55°C. A corresponding improvement in protease resistance was also reported in MB-1Trp (31).

Stabilization can also be achieved by salt bridge formation. It has been shown by Sanchez-Riuz and Makhatadze that there was a higher occurrence of ion pair networks in thermophilic proteins compared to their mesophilic counterpart (32). MB-1 has been designed without any attention for electrostatic contributions, however. At the time of its creation, there was no clear evidence of the effects of ion pairs at the surface of proteins (19).

In this study, we attempt to stabilize MB-1Trp by using both strategies simultaneously: 1- surface hydrophobic residue removal and 2- surface salt bridge formation. We report here the expression and detailed characterization of two single point mutant proteins (Leu19Glu and Leu68Glu), along with the double mutant, of the nutritious protein MB-1Trp using circular dichroism, Trp and ANSA fluorescence, thermal unfolding and protease resistance experiments. Our purpose is to verify the possibility of improving thermal and proteolytic resistance of nutritious *de novo* designed proteins by optimization of surface hydrophilicity.

## MATERIALS AND METHODS

**Preparation of the New Mutants.** Substitution of Leu in position 19 and 68 by Glu on the template protein MB-1Trp (31) was performed using "Altered Sites II mutagenesis kit" (Promega) and "Site-directed Mutagenesis kit" (Stratagene), respectively. The mutagenic oligonucleotides are shown below with the corresponding MB-1Trp sequence (oligo 1 for position 19 and oligo 2 and 3 for position 68). They were purchased from Invitrogen Life Technologies, purified using polyacrylamide gel electrophoresis (PAGE) and phosphorylated.

MB-1Trp: 5'-TTT AAA ACT ATG CAG CTG TTG ACC AAG TCG -3'

Oligo 1: Leu19Glu : 5'-TTT AAA ACT ATG CAG GAA TTG ACC AAG TCG -3'

MB-1Trp: 5'-CGTGG TTC AAA ACG ATG CAG TTG TTA ACG AAG ACC GAG CCCTC-3'

Oligo 2: Leu68Glu: 5'-CGTGG TTC AAA ACG ATG CAG GAG TTA ACG AAG ACC GAG CCCTC-3'

and the reverse complement of Oligo2

Oligo 3: Leu68Glu: 5'-GAGGG CTC GGT CTT CGT TAA CTC CTG CAT CGT TTT GAA CCACG -3'

The mutations were confirmed by dideoxynucleotide DNA sequencing. The combination of both mutations was performed by applying the same method as the one used for Leu68Glu substitution, but on the template MB-1Trp/Leu19Glu. The mutated genes were expressed in pMal-c2 expression vector (New England Biolabs) in fusion with the *MalE* gene. The poly-histidine tag (His-tag) was ligated afterward to the gene by

constructing a chimeric plasmid from the original pMal-c2 MB-1Trp and pMal-c2 MB-1His, which is a similar plasmid containing MB-1 gene with a His-tag at the C-terminal (33). All constructions were confirmed by dideoxynucleotide DNA sequencing.

**Protein Expression and Purification.** All mutants were expressed and purified as described previously (31) with the following modifications. The secondary chromatography involved metal affinity purification, as described in Grundy and coworkers (33). However, ethylenediaminetetraacetate (EDTA) and ethylene glycol-bis( $\beta$ -aminoethylether)-N,N,N',N'-tetraacetate (EGTA) were omitted from the protease inhibitor cocktail to avoid column striping.

**Protein Quantification and Electrophoresis.** Protein concentration was determined by bicinchoninic acid (BCA) assay (Sigma) using bovine serum albumin as a standard. The proteins were visualized by SDS-PAGE using 16% acrylamide-tricine gels and silver staining (Amersham Biosciences) prior to measurements to confirm protein purity.

**Conformational Investigation by Circular Dichroism (CD).** Unless specified otherwise, experiments were conducted at a protein concentration of 0.6 mg/ml with a phosphate buffer (128 mM  $\text{NaH}_2\text{PO}_4$ , pH 6.8). CD measurements were carried out as described previously (31).

**Thermal Denaturation.** The samples were prepared as described in the preceding section. To measure thermostability, temperature was increased from 5 to 98 °C at a rate of 20 °C/h using a Neslab RTE-111 water bath controlled by the Jasco spectropolarimeter software. CD spectra were collected by a Jasco spectropolarimeter J-

720 at every 5 °C from 190 to 260 nm, at a speed of 20 nm/min, and CD signals at 222 nm were collected at every 1 °C.

Thermal stability was calculated assuming a unimolecular, two-state process as previously described (34). The CD signal at 222nm at various temperature was used as the property ( $y$ ) indicative of the extent of unfolding. In the folded state, the parameter  $y = y_f$  and the fraction of folded protein  $f_f$  is equal to 1. When the protein is unfolded, the parameter  $y = y_u$ , and the fraction of unfolded protein  $f_u$  is equal to 1. For intermediate states,  $y$  is given by  $y_f f_f + y_u f_u$ . Thus, by measuring  $y$ , we can calculate the fraction of unfolded protein:  $f_u = (y_f - y)/(y_f - y_u)$ . The equilibrium constant for the unfolding process is  $K_u = f_u/(1-f_u)$  and melting temperatures ( $T_m$ ) are obtained at  $K_u = 1$  (34).

#### **8-Anilino-1-naphthalene sulfonic acid (ANSA) Fluorescence Enhancement.**

Protein concentration was adjusted to 0.1 mg/ml in borate-phosphate buffer (55 mM  $\text{Na}_2\text{B}_4\text{O}_7$ , 55 mM  $\text{NaH}_2\text{PO}_4$ , pH 6.8) and equilibrated 1 hour at room temperature. ANSA was added to a final concentration of 10  $\mu\text{M}$  and equilibrated 5 minutes prior to measurements. Spectra were recorded using an LS50-B Perkin-Elmer fluorometer with an excitation wavelength of 380 nm. Spectra were collected from 410 to 550 nm. Correction for buffer signal and ANSA was applied. Enhancement factors were calculated with the ratio of fluorescence intensity at 480nm of free vs bound ANSA.

**Intrinsic Fluorescence Measurements.** Protein concentration was adjusted to 0.1 mg/ml in borate-phosphate buffer and equilibrated 1 hour at room temperature. Chemical denaturation was performed by adding urea to a final concentration of 8 M and equilibrated at room temperature 5 min before measurements. Spectra were recorded

using an LS50-B Perkin-Elmer fluorometer with an excitation wavelength of 280 nm. Spectra were collected from 290 to 360 nm and corrected for buffer emission.

**Proteolytic Degradation Measurements.** Measurements were done as described previously (31). Cytochrome c (ICN Biomedicals #101467) was used as an internal reference in order to allow for comparison of different digestion experiments.



## RESULTS

**Design Strategy.** In the canonical alpha helix coiled coils folds, heptad positions b, c and f are solvent exposed (35). Hydrophobic side-chains at such exposed positions in native protein are known to force the formation of a hydration structure around them, which contribute to decrease the system entropy and destabilize the native form of the protein (36, 37).

In MB-1, two hydrophobic side chains occupy heptad positions c, which are predicted to be exposed if folded as per design: Leu19 and Leu68. Such residues could promote the previously described destabilizing effect on the protein fold, thus, changing such residues to more polar ones should lead to an increase in stability. Moreover, the substitution of these Leu by acidic residues could lead to the formation of solvent exposed salt bridges with the two nearby Lys residues at the surface of MB-1 as shown in the figure 2.1. Among the acidic residues, Glu is the closest polar isostere of Leu in the naturally occurring amino acids, which approximately limits the effect of the substitution to a change in polarity of the residue's side-chain. Glu has also a higher helix propensity (38) and is a better salt-bridge former than Asp, which is known, in some context, to behave as a helix breaker (39). Helix forming propensity of Glu in both positions was predicted using various criteria which suggest that Glu insertion will promote helix formation in both helices (40, 41). For all the above mentioned reasons, Glu was chosen for the replacement of Leu in both positions 19 and 68.

**Secondary Structure Analysis.** Secondary structures were monitored using far-UV circular dichroism (Figure 2.2). All three mutant proteins Leu68Glu, Leu19Glu, and

double mutant Leu19Glu/Leu68Glu showed the typical signature of helical proteins similar to MB-1Trp, with a slightly lower signal. Using CDSSTR algorithm (42), all mutants were found to have about 40% of their amide groups in a helical environment (Table 1). This percentage is rather low compared to almost all previously described MB-1 mutant proteins shown to have a helical content of about 50 % (19, 31, 43, 44). The loss in helical structure of all mutants is correlated to a comparable gain in  $\beta$ -structure.

**Tertiary Structure Analysis.** Intrinsic fluorescence measurements were performed in order to monitor the environment of the sole Trp residue predicted to be in the core of the protein. As seen in Figure 2.3, Leu68Glu fluorescence spectrum peaked at 335nm with a significant shift upon chemical denaturation to 345nm, which is reminiscent of MB-1Trp. These maxima suggest that Trp62 is protected in the folded state, and become solvated after denaturation. At variance, mutants Leu19Glu and Leu19Glu/Leu68Glu fluorescence spectra peaked at 345nm with no shift upon chemical denaturation (data not shown), indicating a total solvent exposure of the Trp side-chain in both proteins, in either native or denaturing conditions. The mutation Leu19Glu lead to an important modification of the tertiary structure, regardless of the mutation in position 68.

ANSA binding experiments were performed in order to further explore the tertiary structure of the mutant proteins. As seen in Table 2, an enhancement factor around 3 for all proteins was found (same as for the parent protein MB-1Trp) and lower than the value expected for poorly folded proteins (45).

**Effect of the Mutations on Conformational Stability.** Conformational stability was assessed by monitoring the circular dichroism signal at different temperature.

Denaturation curves are shown in Figure 2.4 with the corresponding melting temperatures in Table 1. Mutations Leu19Glu and Leu19Glu/Leu68Glu seem to destabilize the parent molecule MB-1Trp, leading to a loss of 8.4 and 12.5°C in melting temperature respectively. However, the mutation Leu68Glu stabilized MB-1Trp by more than 6°C, an improvement that makes it the most thermoresistant mutant so far in the MB-1 family. Interestingly, Leu68Glu appears folded from 15 up to 40°C. This behavior could lead to a great improvement in proteolytic resistance in cow rumen, in which the temperature is around 39°C. Addition of 500 mM NaCl was performed in order to investigate the contribution of surface electrostatic interactions in mutants Leu68Glu and Leu19Glu. Masking charges in Leu68Glu lead to an important destabilization in agreement with the formation of a stabilizing salt bridge in this mutant. At variance, the presence of salt had no impact on Leu19Glu melting temperature, suggesting that stabilizing salt bridges between Glu19 and neighboring Lys were not formed.

**Effect of the Mutations on Proteolytic Degradation.** The proteolytic resistance of the three mutants was measured and compared to the parent protein MB-1Trp. As can be seen in Figure 2.5, the Leu68Glu mutant resistance to degradation compares with that of MB-1Trp (the parent molecule), and of cytochrome c (a natural protein of similar size). Mutation Leu19Glu led to a destabilization of MB-1Trp, regardless of mutation in position 68. The results suggest a correlation between conformational thermostability and proteolytic resistance for all MB-1Trp proteins analyzed here.

## DISCUSSION

A number of strategies have been investigated and successfully applied for the stabilization of the nutritive protein MB-1: engineering loop motifs (46, 47); insertion of extra- and intra-molecular disulfide bridges (43, 48); polymerization (48); and increasing core hydrophobicity (31). In all cases, increase in thermostability and/or resistance to proteases was achieved in such evolved proteins. Another important feature that has been associated with folding stabilization is the salt bridge. Careful examination of Figure 2.1 reveals that no salt bridges should exist in MB-1 if folded as per design. Reasons for this omission at the time of design have been debated earlier (19, 49), and include 1- the need for a high content in EAA and 2- an attempt to achieve a protein that would be more natural regarding its charged residue composition than existing *de novo* design alpha helical bundles.

Electrostatic interactions have been previously reported to have an important contribution on *de novo* designed proteins native-like folding (50-53). Therefore, we attempted here to achieve better protein properties by inserting negatively charged residues in position 19 and 68, where bridging with neighboring Lys residues at position  $i - 4$  and  $i + 3$  is possible.

The mutation Leu68Glu was found to promote improvements in MB-1Trp characteristics, as supported by fluorescence measurements, thermal denaturation and proteolytic degradation experiments. Although important, similar stabilisation has been observed for surface single point mutation in other studies (54).

Contribution to stability of this mutant protein could come from two factors: 1- removal of surface hydrophobic side-chain (Leu68), and/or 2- formation of stabilizing

salt bridge. The impact of salt on the mutant thermostability strongly suggests that a stabilizing bridge is formed indeed. Note that the stabilisation observed here is not the result of the neutralisation of the helix macrodipole, which requires insertion of negatively charged residues at the N-terminal end of helices.

The mutation Leu19Glu led to destabilization of MB-1Trp protein. Adding salt had no effect on thermal denaturation, suggesting that there were no stabilizing contributions from salt bridges. More importantly, the mutation had a drastic impact on tertiary fold, leading to apparent Trp62 exposure and loss of thermostability. Surprisingly, while two positive modifications at the surface of the mutant Leu19Glu were expected (removal of a non-polar Leu from its surface and formation of stabilizing electrostatic interactions), negative results were recorded, suggesting a major, destabilizing impact on the protein fold. It seems that the structural destabilization provoked by Leu19Glu mutation cannot be corrected by the other mutation Leu68Glu.

Several studies had demonstrated that, in some particular *exposed* environments, polar-to-non-polar side-chain substitution led to stabilisation of the protein (36, 55, 56). Such intuitively destabilizing mutations can in fact help specify a unique, "native" conformation, by restricting access to competing, isoenergetic conformation that would lead to polydispersity, and loss of native traits.

Effect of salt bridges at solvent-exposed sites on proteins is still matter of controversy (57, 58). The topic has been recently discussed by Kumar and Nussinov: it appears that the impact of exposure, solvation and respective positions of the bridge partners on the (de)stabilisation afforded by a mutation complicates the task of protein engineers (59). The different behaviour of our two single point mutants is an example of

the difficulties associated with salt bridge design. The sites chosen for the mutation are expected to be equivalent and the consequences of the mutations also. All data gathered so far on MB-1 and its numerous mutants are compatible with a properly folded alpha-helical bundle (49). In this respect, positions 19 and 68 are expected to be in two equivalent helices (helices I and III). Further, secondary structure prediction and secondary structure propensities indicate that the segments around positions 19 and 68 are equivalent and should be fully helical before and after the mutation Leu to Glu (not shown). Despite this, the mutations have clearly different impacts on MB-1 Trp.

The complexity of salt bridge engineering with regards to energetic contribution to stability of a number of factors has been described by Phelan et al. (60) and Takano et al. (61). Among the factors cited, one finds 1- the favourable coulombic attraction and H-bond formation, 2- the unfavourable cost of side-chain desolvation and fixation, and 3- the unpredictable other contributions, such as global structural modifications and changes in the hydrophobic effect. The mutations studied here are equivalent with regards to item 2. Further, the change in hydrophobic effect is expected to be equivalent. The important modification of the tertiary structure and the absence of salt bridge in the mutants with the mutation Leu19Glu are the reasons for the differences reported here. Such a conclusion suggests another interpretation for the data presented here: the existence of structural differences between position 19 and 68 that probably existed *before* mutagenesis. An investigation is under way that will address the exact location of the helical segments in the proteins studied here, thus allowing to understand the differences between the environments of Leu19 and Leu68.

**ACKNOWLEDGMENT**

The authors thank M.François O. McDuff for artwork. This work was supported by grants from NSERC (Canada) and CORPAQ (Québec) awarded to M.B.

## REFERENCES.

1. FORANO, E. & FLINT, H.J. (2000) Genetically modified organisms: consequences for ruminant health and nutrition. *Ann. Zootech.* **49**, 255–271.
2. WILLIAMSON, M. (1996) Can the risks from transgenic crop plants be estimated? *Trends Biotechnol.* **14**, 449-450.
3. EALING, P.M., HANCOCK, K.R. & WHITE, D.W. (1994) Expression of the pea albumin 1 gene in transgenic white clover and tobacco. *Transgenic Res.* **3**, 344-354.
4. DYER, J.M., NELSON, J.W. & MURAI, N. (1993) Strategies for selecting mutation sites for methionine enhancement in the bean seed storage protein phaseolin. *J. Protein Chem.* **12**, 545-560.
5. GALILI, G. & HOFGEN, R. (2002) Metabolic engineering of amino acids and storage proteins in plants. *Metab. Eng.* **4**, 3-11.
6. ROESLER, K.R. & RAO, A.G. (2001) Rapid gastric fluid digestion and biochemical characterization of engineered proteins enriched in essential amino acids. *J. Agric. Food Chem.* **49**, 3443-3451.
7. CHRISTIANSEN, P., GIBSON, J.M., MOORE, A., PEDERSEN, C., TABE, L. & LARKIN, P.J. (2000) Transgenic *Trifolium repens* with foliage accumulating the high sulphur protein, sunflower seed albumin. *Transgenic Res.* **9**, 103-113.
8. BEACH, L. & TARCZYNSKI, M.C. (2000) Methods of increasing accumulation of essential amino acids in seeds. US Patent 6,127,600.



9. JUNG, R., HASTINGS, C., COUGHLAN, S. & HU, D. (1998) Alteration of amino acids compositions in seeds. U.S. Patent 5,850,016.
10. GASSNER, N.C., BAASE, W.A., MOOERS, B.H. et al. (2003) Multiple methionine substitutions are tolerated in T4 lysozyme and have coupled effects on folding and stability. *Biophys. Chem.* **100**, 325-340.
11. SIGALOV, A.B. & STERN, L.J. (2001) Oxidation of methionine residues affects the structure and stability of apolipoprotein A-I in reconstituted high density lipoprotein particles. *Chem. Phys. Lipids* **113**, 133-146.
12. LAFITTE, D., TSVETKOV, P.O., DEVRED, F. et al. (2002) Cation binding mode of fully oxidised calmodulin explained by the unfolding of the apstate. *Biochim. Biophys. Acta* **1600**, 105-110.
13. DAHIYAT, B.I. & MAYO, S.L. (1997) De novo protein design: fully automated sequence selection. *Science* **278**, 82-87.
14. ROJAS, N.R., KAMTEKAR, S., SIMONS, C.T. et al. (1997) De novo heme proteins from designed combinatorial libraries. *Protein Sci.* **6**, 2512-2524.
15. DALAL, S., BALASUBRAMANIAN, S. & REGAN, L. (1997) Protein alchemy: changing beta-sheet into alpha-helix. *Nat. Struct. Biol.* **4**, 548-552.
16. HILL, R.B. & DEGRADO, W.F. (1998) Solution structure of  $\alpha_2$ D, a native-like de novo designed protein. *J. Am. Chem. Soc.* **120**, 1138-1145.
17. BEAUREGARD, M., GORAJ, K., GOFFIN, V. et al. (1991) Spectroscopic investigation of structure in octarellin (a de novo protein designed to adopt the alpha/beta-barrel packing). *Protein Eng.* **4**, 745-749.

18. BEAUREGARD, M., HEFFORD, M.A. & TEATHER, R.M. (1994) Use of a monoclonal antibody to  $\beta$ -galactosidase for the detection of  $\alpha$ -peptide fusions. *BioTechniques* **16**, 831-838.
19. BEAUREGARD, M., TEATHER, R.M. & HEFFORD, M.A. (1995) Design, expression and initial characterization of MB-1, a de novo protein enriched in essential amino acids. *Biotechnology (N Y)* **13**, 974-981.
20. MACCALLUM, J.D., HEFFORD, M.A., OMAR, S. & BEAUREGARD, M. (1997) Prediction of folding stability and degradability of the de novo designed protein MB-1 in cow rumen. *Appl. Biochem. Biotechnol.* **66**, 83-93.
21. HUANG, X.L., CATIGNANI, G.L. & SWAISGOOD, H.E. (1994) Relative structural stabilities of  $\beta$ -lactoglobulins A and B as determined by proteolytic susceptibility and differential scanning calorimetry. *J. Agric. Food. Chem.* **42**, 1276-1280.
22. LIAO, H.H. (1993) Thermostable mutants of kanamycin nucleotidyltransferase are also more stable to proteinase K, urea, detergents, and water-miscible organic solvents. *Enzyme Microb. Technol.* **15**, 286-292.
23. PARSELL, D.A. & SAUER, R.T. (1989) The structural stability of a protein is an important determinant of its proteolytic susceptibility in *Escherichia coli*. *J. Biol. Chem.* **264**, 7590-7595.
24. KAUZMANN, W. (1958) Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.* **14**, 1-63.

25. WU, L.C. & KIM, P.S. (1997) Hydrophobic sequence minimization of the alpha-lactalbumin molten globule. *Proc. Natl. Acad. Sci. U S A.* **94**, 14314-14319.
26. CORDES, M.H., DAVIDSON, A.R. & SAUER, R.T. (1996) Sequence space, folding and protein design. *Curr. Opin. Struct. Biol.* **6**, 3-10.
27. PONDER, J.W. & RICHARDS, F.M. (1987) Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes. *J. Mol. Biol.* **193**, 775-791.
28. OHMURA, T., UEDA, T., OOTSUKA, K., SAITO, M. & IMOTO, T. (2001) Stabilization of hen egg white lysozyme by a cavity-filling mutation. *Protein Sci.* **10**, 313-320.
29. GOLOVANOV, A.P., VOLOVANOV, G. & ARSENEV, A.S. (2000) Stabilization of proteins by enhancement of inter-residue hydrophobic contacts: lessons of T4 lysozyme and barnase. *J. Biomol. Struct. Dyn.* **18**, 477-491.
30. CHOE, S.E., LI, L., MATSUDAIRA, P.T., WAGNER, G. & SHAKHNOVICH, E.I. (2000) Differential stabilization of two hydrophobic cores in the transition state of the villin 14T folding reaction. *J. Mol. Biol.* **304**, 99-115.
31. GAGNON, M.C., WILLIAMS, M., DOUCET, A. & BEAUREGARD, M. (2000) Replacement of tyr62 by trp in the designer protein milk bundle-1 results in significant improvement of conformational stability. *FEBS Lett.* **484**, 144-148.
32. SANCHEZ-RUIZ, J.M. & MAKHATADZE, G.I. (2001) To charge or not to charge? *Trends Biotechnol.* **19**, 132-135.

33. GRUNDY, J.E., WIRTANEN, L.Y. & BEAUREGARD, M. (1998) Addition of a poly-(6X) His tag to Milk Bundle-1 and purification using immobilized metal-affinity chromatography. *Protein Expr. Purif.* **13**, 61-66.
34. PACE, C.N., SHIRLEY, B.A. & THOMSON, J.A. (1989) Measuring the conformational stability of a protein. In *Protein Structure - A Practical Approach*. (Creighton, T.E., ed.) IRL press, Oxford, pp-311-330.
35. HODGES, R.S., SODEK, J., SMILLIE, L.B. & JURASEK, L. (1972) Tropomyosin: Amino acid sequence and coiled-coil structure. *Cold Spring Harb. Symp. Quant. Biol.* **37**, 299-310.
36. FUNAHASHI, J., TAKANO, K., YAMAGATA, Y. & YUTANI, K. (2000) Role of surface hydrophobic residues in the conformational stability of human lysozyme at three different positions. *Biochemistry* **39**, 14448-14456.
37. SAENGER, W. (1987) Structure and dynamics of water surrounding biomolecules. *Annu. Rev. Biophys. Biophys. Chem.* **16**, 93-114.
38. PACE, C.N. & SCHOLTZ, J.M. (1998) A helix propensity scale based on experimental studies of peptides and proteins. *Biophys. J.* **75**, 422-427.
39. BURKHARD, P., IVANINSKII, S. & LUSTIG, A. (2002) Improving coiled-coil stability by optimizing ionic interactions. *J. Mol. Biol.* **318**, 901-910.
40. GARNIER, J., GIBRAT, J.F. & ROBSON, B. (1996) GOR method for predicting protein secondary structure from amino acid sequence. *Methods Enzymol.* **266**, 540-553.

41. KOEHL, P. & LEVITT, M. (1999) Structure-based conformational preferences of amino acids. *Proc. Natl. Acad. Sci. USA* **96**, 12524-12529.
42. JOHNSON, W.C. (1999) Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins* **35**, 307-312.
43. DOUCET, A., WILLIAMS, M., GAGNON, M.C., SASSEVILLE, M. & BEAUREGARD, M. (2002) Engineering nutritious proteins: improvement of stability in the designer protein MB-1 via introduction of disulfide bridges. *J. Agric. Food Chem.* **50**, 92-98.
44. WILLIAMS, M., GAGNON, M.C., DOUCET, A. & BEAUREGARD, M. (2002) Design of high essential amino acid proteins: two design strategies for improving protease resistance of the nutritious MB-1 protein. *J. Biotechnol.* **94**, 245-254.
45. SEMISOTNOV, G.V., RODIONOVA, N.A., RAZGULYAEV, O.I., UVERSKY, V.N., GRIPAS', A.F. & GILMANSHIN, R.I. (1991) Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers* **31**, 119-128.
46. PARKER, M.H. & HEFFORD, M.A. (1997) A consensus residue analysis of loop and helix-capping residues in four-alpha-helical-bundle proteins. *Protein Eng.* **10**, 487-496.
47. PARKER, M.H. & HEFFORD, M.A. (1998) Introduction of potential helix-capping residues into an engineered helical protein. *Biotechnol. Appl. Biochem.* **28**, 69-76.
48. MORRISON, J.J., McNALLY, J.D., NAVIDZADEH, A. & BEAUREGARD, M. (2000) Development of an optimized feeding technology for dairy cows: improvement in resistance to ruminal proteases in the de novo-designed protein MB-1. *Appl. Biochem. Biotechnol.* **87**, 247-264.

49. HEFFORD, M.A., DUPONT, C., MACCALLUM, J., PARKER, M.H. & BEAUREGARD, M. (1999) Characterization of MB-1: A dimeric helical protein with a compact core. *Eur. J. Biochemistry* **262**, 467-474.
50. MEIER, M., LUSTIG, A., AEBI, U. & BURKHARD, P. (2002) Removing an interhelical salt bridge abolishes coiled-coil formation in a de novo designed peptide. *J. Struct. Biol.* **137**, 65-72.
51. JOHANSSON, J.S., GIBNEY, B.R., SKALICKY, J.J., WAND, A.J. & DUTTON, P.L. (1998) A native-like three- $\alpha$ -helix bundle protein from structure-based redesign: a novel maquette scaffold. *J. Am. Chem. Soc.* **120**, 3881-3886.
52. JIANG, X., BISHOP, E.J. & FARID, R.S. (1997) A de novo designed protein with properties that characterize natural hyperthermophilic proteins. *J. Am. Chem. Soc.* **119**, 838-839.
53. BETZ, S.F., LIEBMAN, P.A. & DEGRADO, W.F. (1997) De novo design of native proteins: characterization of proteins intended to fold into antiparallel, rop-like, four-helix bundles. *Biochemistry* **36**, 2450-2458.
54. SANZ-APARICIO, J., HERMOSO, J.A., MARTINEZ-RIPOLL, M., GONZALEZ, B., LOPEZ-CAMACHO, C. & POLAINA, J. (1998) Structural basis of increased resistance to thermal denaturation induced by single amino acid substitution in the sequence of beta-glucosidase A from *Bacillus polymyxa*. *Proteins* **33**, 567-576.

55. POSO, D., SESSIONS, R.B., LORCH, M. & CLARKE, A.R. (2000) Progressive stabilization of intermediate and transition states in protein folding reactions by introducing surface hydrophobic residues. *J. Biol. Chem.* **275**, 35723-35726.
56. SCHWEHM, J.M., KRISTYANNE, E.S., BIGGERS, C.C. & STITES, W.E. (1998) Stability effects of increasing the hydrophobicity of solvent-exposed side chains in staphylococcal nuclease. *Biochemistry* **37**, 6939-6948.
57. STROP, P. & MAYO, S.L. (2000) Contribution of surface salt bridges to protein stability. *Biochemistry* **39**, 1251-1255.
58. IBARRA-MOLERO, B., LOLADZE, V.V., MAKHATADZE, G.I. & SANCHEZ-RUIZ, J.M. (1999) Thermal versus guanidine-induced unfolding of ubiquitin. An analysis in terms of the contributions from charge-charge interactions to protein stability. *Biochemistry* **38**, 8138-8149.
59. KUMAR, S. & NUSSINOV, R. (2002) Close-range electrostatic interactions in proteins. *Chembiochem.* **3**, 604-617.
60. PHELAN, P., GORFE, A.A., JELESAROV, I., MARTI, D.N., WARWICKER, J. & BOSSHARD, H.R. (2002) Salt bridges destabilize a Leu zipper designed for maximized ion pairing between helices. *Biochemistry* **41**, 2998-3008.
61. TAKANO, K., TSUCHIMORI, K., YAMAGATA, Y. & YUTANI, K. (2000) Contribution of salt bridges near the surface of a protein to the conformational stability. *Biochemistry* **39**, 12375-12381.

Table 1 : Secondary structure analysis for MB-1Trp and its variants

Proteins	Helix $\pm 3\%$	$\beta$ -Strand $\pm 3\%$	$\beta$ -Turn $\pm 3\%$	Others $\pm 3\%$	Tm ( $^{\circ}\text{C}$ ) $\pm 1^{\circ}\text{C}$	Tm ( $^{\circ}\text{C}$ ) with salt $\pm 1^{\circ}\text{C}$
MB-1 Trp	51	8	14	28	55.0	-
Leu19Glu	37	15	13	35	47.6	47.1
Leu68Glu	37	18	12	33	61.1	46.6
Leu19Glu/Leu68Glu	38	14	12	36	42.5	-



Table 2 : Tryptophane and ANSA fluorescence experiments of MB-1 Trp and its variants

Proteins	$\lambda_{\max}$ Trp (nm)	Core Fluidity
MB-1Trp	333	3
Leu19Glu	345	2.5
Leu68Glu	335	3.3
Leu19Glu/Leu68Glu	345	3

## FIGURE LEGENDS

Figure 2.1: Predicted location of secondary structures in MB-1Trp. The structure is flattened with the core-forming hydrophobic residues in background. Leu 19 and 68 are dark-circled; 6 His were added for purification purposes.

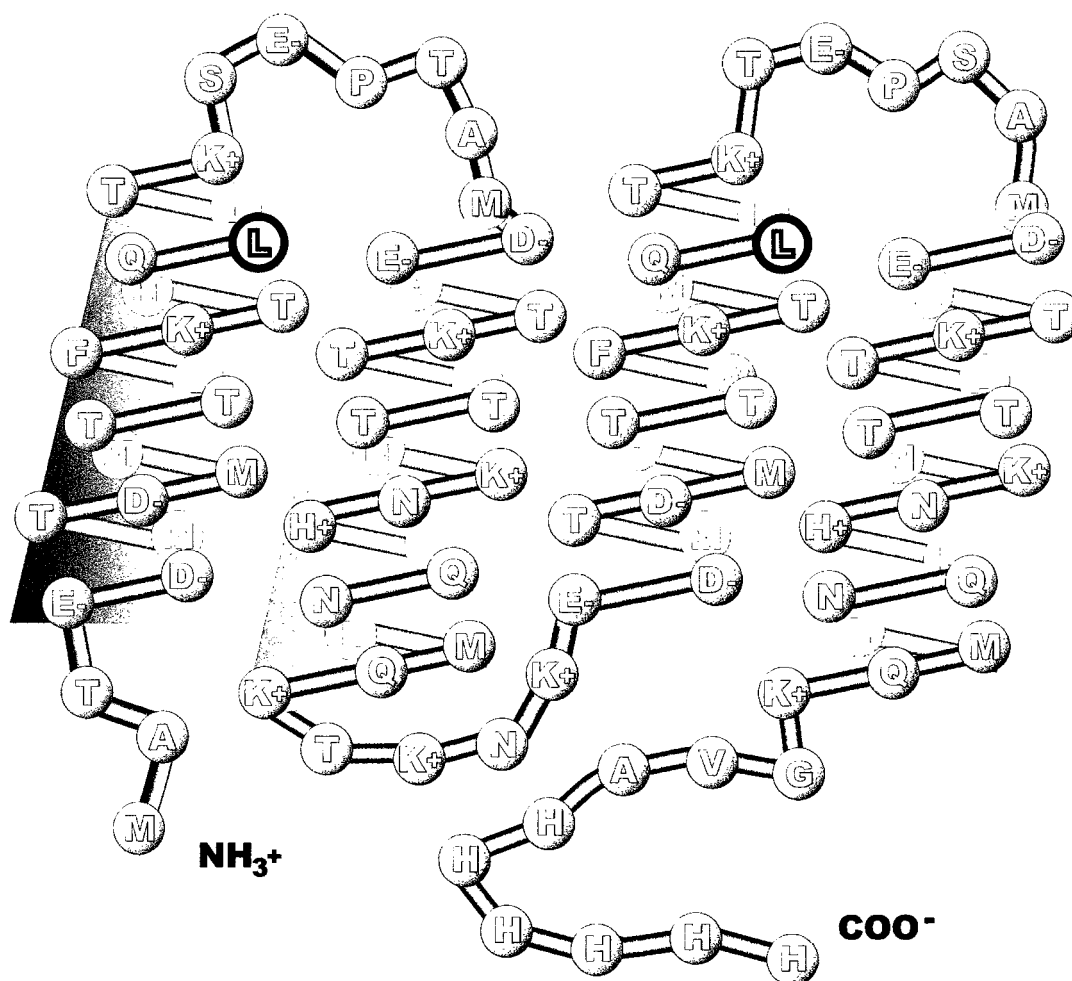
Figure 2.2: CD spectra of MB-1Trp and its variants: (○) MB-1Trp; (□) Leu19Glu; (◇) Leu68Glu; (Δ) Leu19Glu/Leu68Glu.

Figure 2.3: Tryptophan fluorescence spectra of MB-1Trp and its variants. Spectra were normalized to 100 a.u. in order to allow for direct comparison of maxima. Same symbols as above.

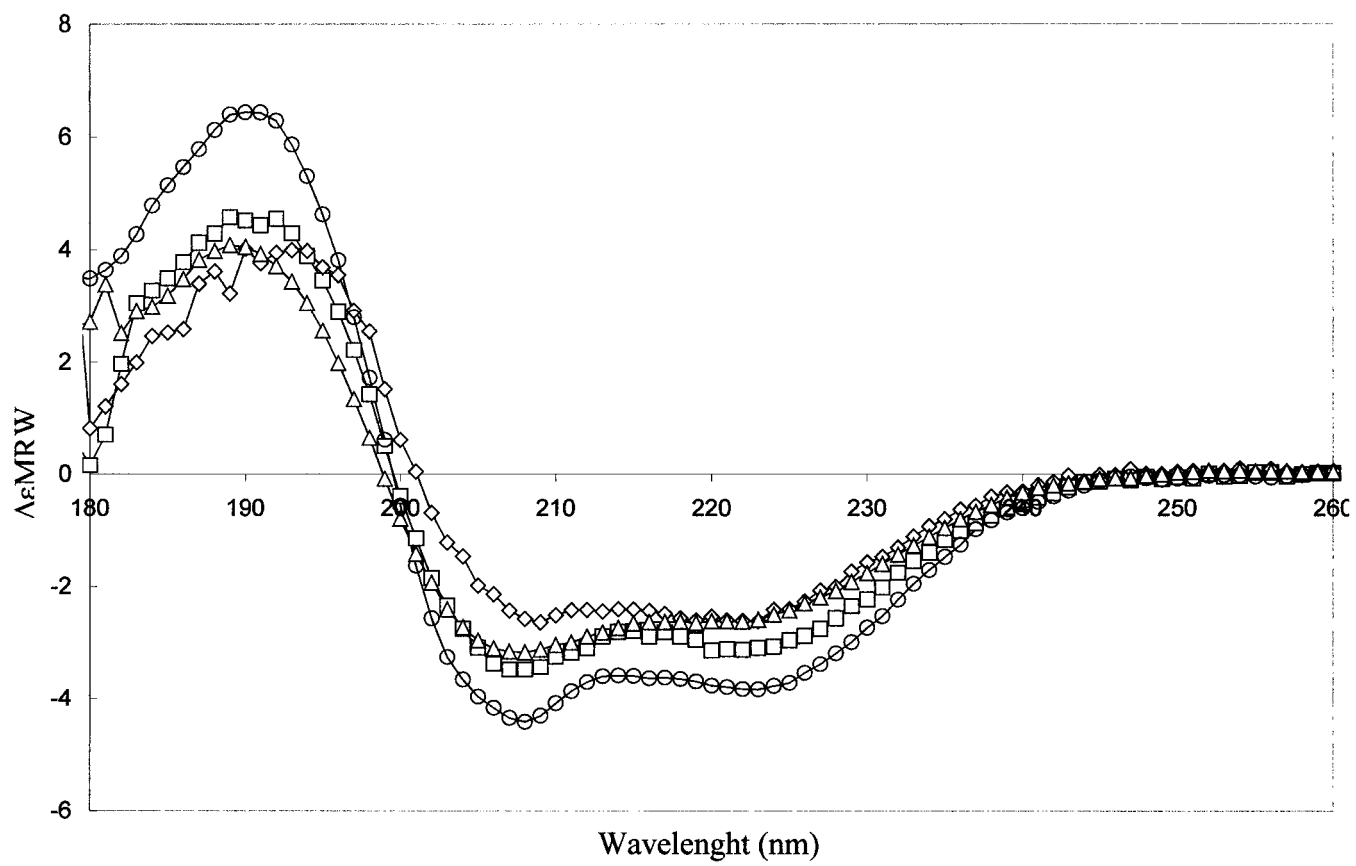
Figure 2.4: Thermal denaturation of MB-1Trp and its variants monitored using molar ellipticity at 222 nm. Symbols are the same as for Figure 2.2.

Figure 2.5: Proteolytic degradation of MB-1Trp and its variants. Percentages of degradation are based on the protein disappearance during incubation with proteases. Degradation curves for cytochrome c (\*), the protein standard used to normalize all digestion experiments, and that of MB-1 (+), the ancestor of MB-1 family, are shown for

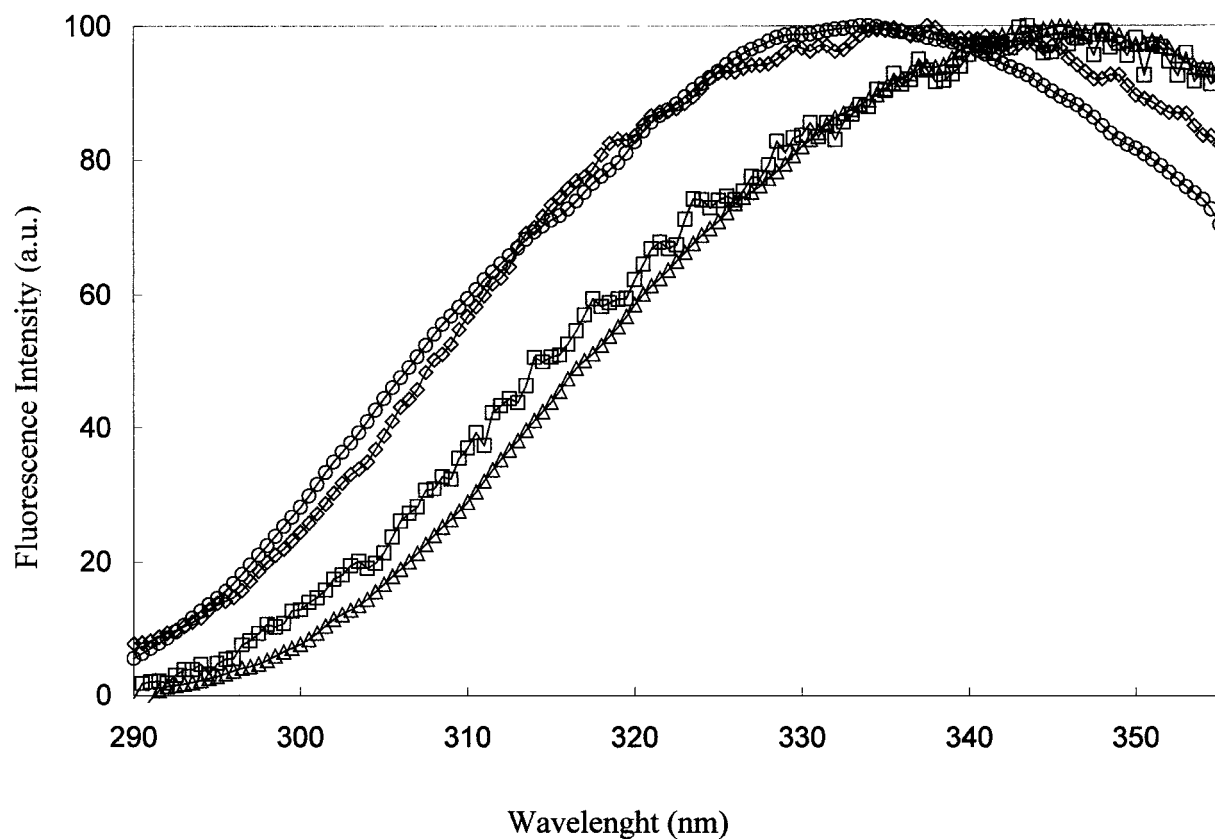
comparison. Other symbols are same as above, Leu19Glu data superimposed those obtained for the double mutant.



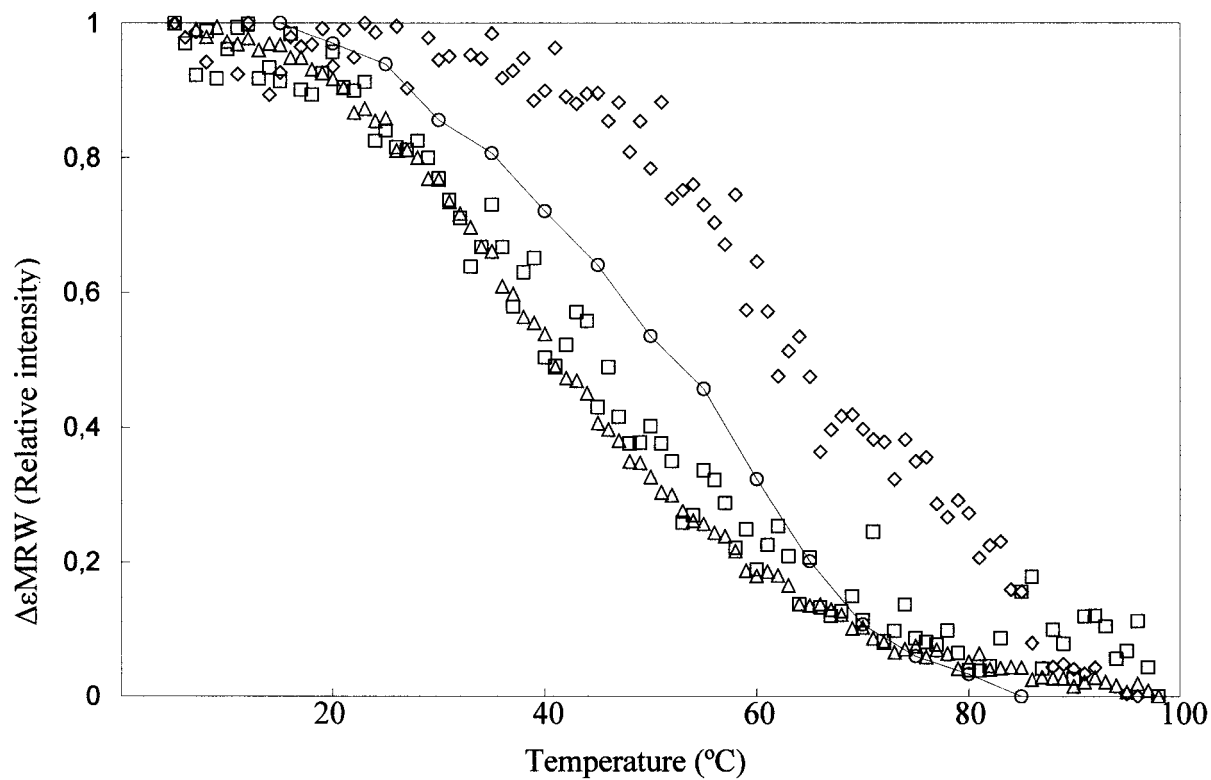
**Figure 2.1: Predicted location of secondary structures of MB-1TrpHis.** The structure is flattened with the core-forming hydrophobic residues in background. Leu 19 and 68 are dark-circled; 6 His were added for purification purposes.



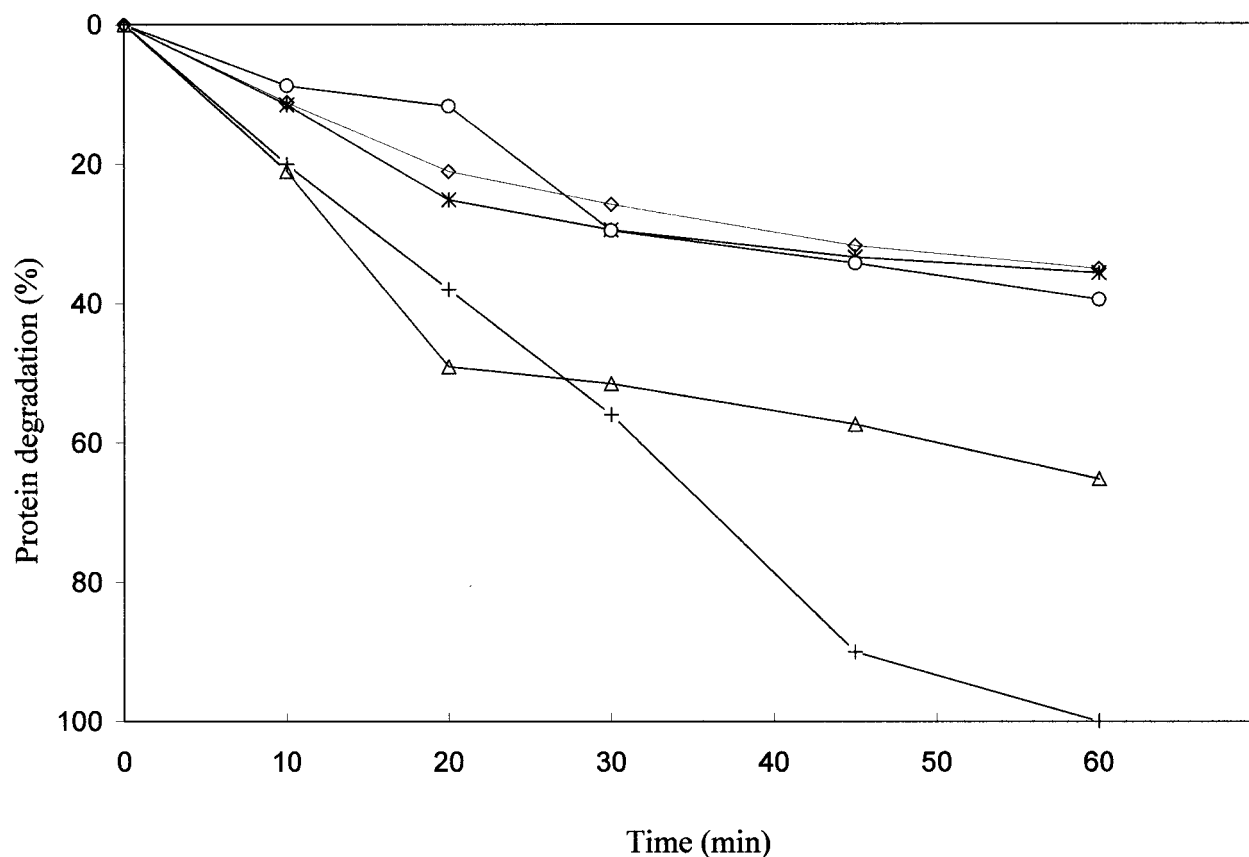
**Figure 2.2:** CD spectra of MB-1Trp and its variants: (○) MB-1Trp; (□) Leu19Glu; (◇) Leu68Glu; (Δ) Leu19Glu/Leu68Glu.



**Figure 2.3: Tryptophane fluorescence spectra of MB-1Trp and its variants.** Spectra were normalized to 100 a.u. in order to allow for direct comparison of maxima. (○) MB-1Trp; (□) Leu19Glu; (◇) Leu68Glu; (Δ) Leu19Glu/Leu68Glu.



**Figure 2.4: Thermal denaturation of MB-1Trp and its variants monitored using molar ellipticity at 222 nm. (○) MB-1Trp; (□) Leu19Glu; (◇) Leu68Glu; (Δ) Leu19Glu/Leu68Glu.**



**Figure 2.5: Proteolytic degradation of MB-1Trp and its variants.** Percentages of degradation are based on the protein disappearance during incubation with proteases. Degradation curves for cytochrome c (\*), the protein standard used to normalize all digestion experiments, and that of MB-1 (+), the ancestor of MB-1 family, are shown for comparison. Other symbols are same as above, Leu19Glu data superimposed those obtained for the double mutant.



## 2.2 Autres résultats

Bien que l'article contienne la majorité des résultats de mon projet, une observation non-mentionnée mérite qu'on s'y attarde.

Dans un premier temps, la chromatographie sur résine d'amylose permet d'isoler, à partir d'un lysat cellulaire, les protéines qui ont une affinité pour le maltose (l'amylose étant un polymère de maltose). Notre protéine, étant fusionnée à la "maltose-binding protein" (MBP), est récupérée lors de l'élution de la colonne d'amylose. Dans un deuxième temps, l'éluat, après une exposition au facteur X<sub>a</sub> (une protéase spécifique au site Ile-Glu-Gly-Arg placé entre notre protéine et la MBP) est injecté sur une résine DEAE sépharose échangeuse d'anions. Historiquement, MB-1Trp a toujours été lavé de la colonne sans gradient de sel tandis que MBP restait accrochée (Gagnon et al. 2000). Cette observation est corroborée par la prédiction de point isoélectrique, soit 6.72 pour MB-1Trp et 5.08 pour MBP, qui prédit une plus grande proportion de charges négatives à pH 8.0 pour la MBP que pour MB-1Trp (Compute pI/Mw tool, Expasy home page, expasy.org). Cependant, nos mutants ont changé de comportement en s'accrochant à la résine DEAE lors du lavage et en s'éluant dans le gradient de sel avec la MBP.

Les mutations aux positions 19 et 68 ajoutent théoriquement deux groupements carboxyles à la surface de la protéine. Ces groupements déplacent le point isoélectrique de la protéine vers des pI acides comme nous le démontre le logiciel de prédiction de point isoélectrique "Compute pI" avec les valeurs 6.08 et 5.63 pour les simples mutants et

le double mutant respectivement. Ces points isoélectriques sont calculés pour les peptides dénaturés et doivent donc être pris à titre indicatif seulement étant donné que le point isoélectrique est déterminé en réalité par la balance des charges de surface d'une protéine.

En plus de nous obliger à opter pour la construction d'un système utilisant la purification par une affinité au nickel comme purification subséquente, ce changement de point isoélectrique nous pousse à croire que nos mutations aux positions 19 et 68 sont bel et bien à la surface de la protéine comme la prédiction de structure secondaire le propose.

## CHAPITRE III

### DISCUSSION

#### 3.1 L'effet des ponts salins sur la stabilité de MB-1Trp

L'effet des ponts salins à la surface des protéines est toujours controversé. Plusieurs auteurs montrent une contribution stabilisante de leurs ponts salins, soit par la déstabilisation d'une protéine suite à leur retrait ou par la stabilisation de la protéine par leur ajout (Merz et al. 1999; Lounnas et al. 1997). D'autres ont montré l'inverse, soit des contributions déstabilisantes suite à l'insertion de ponts salins à la surface de la protéine (Lee et al. 2002; Phelan et al. 2002).

L'ambiguïté de l'effet des ponts salins vient probablement de la complexité des contributions qui entrent en jeu lors de leur formation. L'équipe de Takano et celle de Phelan ont bien cerné les contributions impliquées lors de la formation d'un pont salin exposé. Une version combinée de leurs équations est présentée ci-bas :

$$\text{Équation [1]} \quad \Delta G_{\text{Pont Salin}} = \Delta G_{\text{Attraction}} + \Delta G_{\text{Desolvation}} + \Delta G_{\text{Entropique}} + \Delta G_{\text{Autre}}$$

où le  $\Delta G_{\text{Pont Salin}}$  représente le changement d'énergie libre de la protéine lors de la formation du pont salin; le  $\Delta G_{\text{Attraction}}$ , la contribution coulombique attractive (deux

charges l'une par rapport à l'autre); le  $\Delta G_{\text{Desolvation}}$ , le coût déstabilisant qu'implique le retrait des molécules de solvant des chaînes latérales impliquées dans le pont salin;  $\Delta G_{\text{Entropique}}$ , le coût entropique de fixation des chaînes latérales dans une position précise pour former le pont salin; et le  $\Delta G_{\text{Autre}}$  regroupe les contributions imprédictibles, c'est-à-dire les modifications de structures secondaires et tertiaires et les changements dans l'effet hydrophobe.

La différence d'énergie libre entre l'état natif et l'état dénaturé d'une protéine est relativement faible : les contributions stabilisantes (effet hydrophobe, interactions électrostatiques, pont hydrogène, interactions de Van der Waals) d'une part et les contributions déstabilisantes (entropie conformationnelle, température) d'autre part, se distinguent seulement par quelques  $\text{kJ mol}^{-1}$  pour une protéine native. Il en est de même pour un pont salin, sa formation peut être stabilisante ou non selon le cas. Pour permettre la formation d'un pont salin, la différence d'énergie libre globale ( $\Delta G_{\text{Pont Salin}}$ ) doit être négative (stabilisante). Pour ce faire, les contributions stabilisantes ( $\Delta G_{\text{Attraction}} + \Delta G_{\text{Autre}}$ ) doivent être plus importantes que les contributions déstabilisantes ( $\Delta G_{\text{Desolvation}} + \Delta G_{\text{Entropique}} + \Delta G_{\text{Autre}}$ ).

Kumar et ses collègues ont déterminé que la force attractive coulombique d'un pont salin est directement reliée à l'orientation et à la distance des chaînes latérales l'une par rapport à l'autre (Kumar et al. 2002). Le mauvais alignement ou l'éloignement réduit la contribution d'attraction des charges opposées. Des études faites par résonance

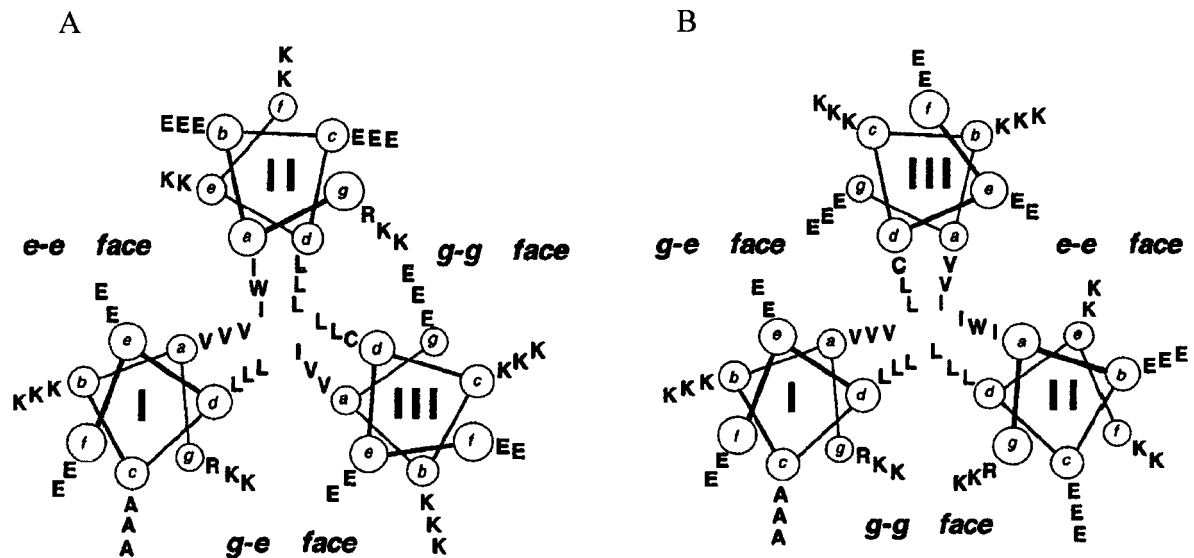
magnétique nucléaire (RMN) avec l'hétérodimère c-Myc-Max ont montré que la force et la géométrie des ponts salins fluctuent énormément avec les différents conformères qu'adopte la protéine. Ces ponts salins s'interconvertissent d'un état stabilisant à déstabilisant. Les ponts salins stabilisants sont ceux qui se retrouveront plus souvent dans des conformations stabilisantes que déstabilisantes. Ce résultat nous rappelle que les ponts salins, tout comme les protéines, sont dynamiques et non pas statiques (Kumar et al. 2000).

L'insertion d'un pont salin à la surface de MB-1Trp s'est avérée une bonne stratégie pour améliorer ses caractéristiques dans le contexte où aucune structure tridimensionnelle n'est disponible. Certains exemples de protéines *de novo* démontrent un réseau de ponts salins très développé. Le baril à trois hélices conçu par Johansson et ses collègues possède, comme on peut le voir à la figure 3.1, un réseau de sept ponts salins qui spécifie la topologie des hélices tout en la stabilisant (Johansson et al. 1998). Il en est de même pour la grande majorité des protéines hélicales *de novo* (Meier et al. 2002; Hill et al. 2000; Jiang et al. 1997; Betz et al. 1997). Notre approche se distingue par le fait que nous visons la formation de ponts salins de surface ponctuels sur notre protéine.

La réussite de notre stratégie est démontrée par les caractéristiques du mutant Leu68Glu. La formation du pont salin a été démontrée par la perte de thermostabilité lors d'expériences de masquage d'interactions électrostatiques par une haute concentration de

sel. La conservation de la conformation et des structures secondaires montre que la mutation n'a pas d'effet inattendu sur la protéine. On peut également constater après la comparaison des courbes de dénaturation que Leu68Glu demeure complètement pliée à 40°C, au-delà de la température du rumen de la vache (donc celle à laquelle la protéine devra résister), tandis que MB-1Trp a déjà perdu 20% de ses structures secondaires à la même température. Aucune augmentation de la coopérativité de la dénaturation n'a été observée dans aucun des mutants, indiquant qu'aucune amélioration n'a été faite à la compaction spécifique du cœur hydrophobe et au comportement natif naturel de la protéine. Cette constatation n'est cependant pas une surprise compte tenu du fait que les mutations devraient affecter principalement la surface de la protéine.

Le comportement du mutant Leu19Glu a été plus surprenant. Ce mutant a subi un changement de structure tertiaire important ayant probablement mené à un changement des voisins du tryptophane, de concert avec une perte de thermostabilité de près d'une dizaine de °C. Le changement de structure tertiaire est d'ailleurs accompagné par l'absence de formation du pont salin. Ce comportement est d'autant plus inattendu que la substitution a lieu dans un contexte qui devrait être homologue, pour ne pas dire identique, à la substitution Leu68Glu. Évidemment, la contribution stabilisante de la mutation Leu68Glu lorsque appliquée sur le variant Leu19Glu n'apparaît pas, probablement dû aux changements dans la structure de la protéine induits par la substitution Leu19Glu qui empêche la formation du pont salin en position 68.



**Figure 3.1: A) Exemple de la protéine en baril à trois hélices ayant un réseau de ponts salins E-K et E-R aux interfaces (positions "g" et "e" sur l'heptade). B) La même protéine montrant la contribution de spécification du repliement des interactions électrostatiques. La topologie B serait probablement déstabilisée par rapport à la A, car la conservation du repliement B aurait un coût en énergie libre plus élevé, dues aux interactions répulsives sur les faces g-e et g-g. Cette stratégie de spécification de la topologie s'appelle "negative design". (Tiré de Johansson et al. 1998)**

### 3.2 Les éléments qui différencient les mutations aux positions 19 et 68

Afin de comprendre ce qui fait que les deux mutations ne donnent pas des résultats semblables, nous allons décortiquer les facteurs en jeu.

#### 3.2.1 Quelles sont les contributions impliquées dans la stabilisation du mutant Leu68Glu et qui devraient être présentes dans le mutant Leu19Glu?

L'équation 1 (p.67) énumère les contributions à considérer pour répondre à cette question. Premièrement, la contribution attractive, due aux charges opposées, semble donner une bonne partie de l'effet observé dans le mutant Leu68Glu, car on observe une diminution de stabilité lors du masquage avec le sel. La substitution donnant cette même possibilité au mutant Leu19Glu est présente, mais ne semble pas contribuer à un effet attractif. En effet, il n'y a aucun changement lors des expériences de dénaturation thermique en présence d'une forte concentration de sel.

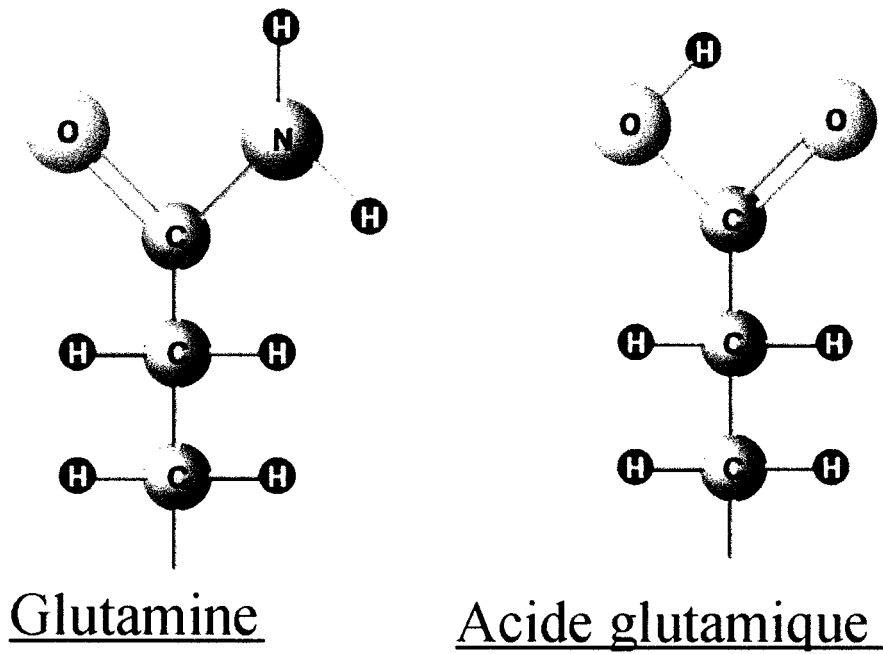
Deuxièmement, le coût déstabilisant de désolvation des chaînes latérales est difficile à déterminer. Cependant, on peut estimer qu'il est moins important que la contribution attractive, car autrement la formation du pont salin serait impossible dans le mutant Leu68Glu. De plus, ce coût devrait être identique pour le mutant Leu19Glu, car les deux résidus sont dans des positions semblables dans leur hélice respective.



Troisièmement, le coût entropique de fixation des chaînes latérales dans une seule conformation permettant la formation du pont salin devrait être théoriquement le même dans les deux mutants si les deux positions sont dans un environnement semblable.

Quatrièmement, il y a les autres contributions à la stabilité. Le changement de structures secondaires, observé dans chacun des mutants, ne semble avoir qu'un lien indirect avec une ou l'autre des mutations. Effectivement, le même changement de structures secondaires a donné des effets différents sur la stabilité des trois mutants. Une hypothèse plausible serait que les cinq ou six derniers acides aminés (positions 94-100) adoptent une conformation  $\beta$  et que la queue de poly-histidines elle-même soit en conformation désordonnée. Cette hypothèse permet d'expliquer les pourcentages de structures secondaires calculées pour les trois protéines. Cette hypothèse est également appuyée par une analyse faite par le logiciel de prédiction de structures secondaires GOR IV (non montrée) qui démontre le potentiel de la queue de poly-histidine à provoquer la transition  $\alpha$ - $\beta$  des quatre derniers acides aminés en C-terminal (97-100).

La contribution de la substitution à l'effet hydrophobe global est aussi difficile à évaluer. La structure prédite nous pousse à croire que les deux mutants devraient avoir un changement d'effet hydrophobe comparable, car elles sont dans des positions similaires. Une expérience intéressante qui permettrait d'isoler la contribution à l'effet hydrophobe serait de faire un autre cycle de mutations doubles, comme celui-ci, mais en utilisant la glutamine comme substituant. Comme on peut le voir à la figure 3.2, les chaînes latérales



**Figure 3.2: La chaîne latérale du glutamate et de la glutamine.** On peut y voir la différence dans les chaînes latérales : le groupement amine des glutamines est remplacé par un groupement hydroxyl dans les glutamates.

du glutamate et de la glutamine sont très semblables avec seulement un groupement amine au lieu d'un groupement hydroxyle sur le carbone  $\gamma$ . Cette expérience montrerait l'effet du retrait de deux groupements méthylènes de la surface de la protéine en remplaçant par un groupement polaire non-chargé. On découplerait la contribution de l'effet hydrophobe de celle des ponts salins. Cependant, rien ne pousse à croire que la contribution à l'effet hydrophobe soit différente pour les substitutions en position 19 et 68.

Il n'y a aucune évidence d'un changement de structure tertiaire provoqué par la mutation Leu68Glu, tel que montré par la fluorescence du tryptophane et de l'ANSA. Cependant, il a été démontré, par les présentes expériences, que la substitution Leu19Glu provoque un changement important de la structure tertiaire et de l'organisation du cœur hydrophobe. Il y a donc une différence fondamentale entre ces deux mutations.

### 3.2.2 Qu'est-ce qui peut expliquer une différence entre ces deux substitutions apparemment semblables (Leu19Glu vs Leu68Glu)?

En utilisant le logiciel GOR IV, la structure des trois mutants a été prédite ayant un contenu élevé en hélice, soit similaires à MB-1Trp, exception faite de l'effet de la queue d'histidines que nous avons traité précédemment. Les prédictions de structures secondaires sont : 70-71% en hélice, 21-23% en coude désordonné et 7-10% en structure  $\beta$ . Les différents changements de caractéristiques observées dans les variants qui sont

présentés ici ne peuvent donc s'expliquer par des changements au niveau de la capacité de la structure primaire à former des hélices.

Au niveau de la structure probable des mutants (figure 1.5), les seules différences entre les deux sites sont les suivant :

- La proximité du site de mutation avec le résidu Trp62;
- La localisation qui est soit dans une hélice terminale, soit dans une hélice intermédiaire.

Intuitivement, c'est autour de la mutation Leu19Glu que l'on doit d'abord chercher une explication. Des exemples de déstabilisation d'hélices terminales similaires au cas de Leu19Glu ont été rapportés dans la littérature (Chu et al. 2002).

Il est également possible que, malgré nos prédictions, la séquence exacte des hélices et la position des coudes ne soient pas telles que prédites. Cette hypothèse est plausible étant donné la différence entre la prédiction de structures secondaires et les proportions de structures secondaires mesurées par dichroïsme circulaire. La position 19 pourrait en réalité être à l'intérieur du cœur hydrophobe. Ainsi, un changement d'une chaîne latérale hydrophobe pour une chaîne polaire pourrait provoquer un changement de l'acide aminé voisin du Trp62 qui expliquerait le changement des caractéristiques de fluorescence du mutant Leu19Glu. Une déstabilisation du cœur hydrophobe qui exposerait le Trp62 au solvant pourrait également provoquer cet effet. Alternativement, le pont salin pourrait également s'être formé, mais la position de l'hélice I projetterait la

paire Glu-Lys à l'intérieur du cœur hydrophobe, créant une paire ionique très forte, insensible à une haute concentration de sel, qui déplacerait la fluorescence du mutant vers le rouge tout en déstabilisant toute la structure tel qu'observé dans le cas de Leu19Glu et du double mutant. Ces hypothèses ne peuvent être ignorées puisque nous n'avons pas d'information sur la structure tridimensionnelle précise pour les protéines de la famille MB-1.

L'effet des résidus hydrophobes et aromatiques à la surface des protéines naturelles a dernièrement été traité d'une façon tout à fait surprenante par quelques équipes de recherche. Le groupe de Tisi et Evens attribue, à ces petits regroupements hydrophobes de surface, un rôle de déterminant de la structure tertiaire, en limitant le nombre de conformères isoénergétiques disponibles pour la protéine. Il souligne la prédominance de l'effet hydrophobe comme élément poussant le repliement des protéines, mais identifie les interactions de surface comme étant des contributions à la structure fine de l'état natif. D'autres groupes identifient également le rôle de la structure d'hydratation de la chaîne latérale hydrophobe comme étant un élément important dans l'effet d'une mutation hydrophobe de surface (Funahashi et al. 2000; Poso et al. 2000; Schwemhe et al. 1998; Tisi et al. 1995). Comme ces informations ne sont disponibles que par la résolution de la structure tridimensionnelle de la protéine, ces éléments ne peuvent qu'être des pistes de recherche pour l'instant. Donc, l'effet d'une mutation de surface est extrêmement dépendant du contexte.

En conclusion, le retrait de chaîne latérale hydrophobe combiné à l'ajout de ponts salins à la surface de la protéine MB-1Trp s'est avéré être une stratégie valable pour augmenter sa thermostabilité. Cette augmentation de stabilité et la conservation de l'état natif jusqu'à 40°C devrait nous assurer un meilleur rendement lors de son application en agro-biotechnologie (meilleure accumulation dans une plante transgénique, meilleure résistance dans le rumen, etc.). Il est important de rappeler que l'augmentation de la stabilité de la protéine MB-1 est capitale, comme pour toute autre protéine appliquée en nutrition des ruminants. Dans un premier temps, l'accumulation d'une protéine idéale dans le cytoplasme d'une plante, étape préalable à sa production, à son stockage et à sa livraison à la vache, est directement liée à sa capacité à résister à la machinerie cellulaire végétale de dégradation protéolytique. Dans un deuxième temps, la protéine idéale devrait être de la plus grande stabilité possible dans une gamme de pH de 5.5 à 7.0. Ce pH est celui du rumen, premier estomac de la vache contenant la flore bactérienne susceptible de dégrader la protéine d'intérêt. Dans un troisième temps, la protéine idéale devrait perdre sa stabilité dans une gamme de pH de 2.0 à 3.0, le pH de l'abomasum, pour permettre sa dégradation et son assimilation (Papas et al. 1997). En ce sens, il n'y a aucune limite à la stabilité de la protéine exprimée tant et aussi longtemps qu'une certaine déstabilisation est possible à bas pH. Dans le contexte de MB-1, l'ajout de ponts salins est une stratégie d'autant plus intéressante, car les charges permettant la formation de l'interaction ionique entre les chaînes latérales de la Lys et du Glu sont présentes qu'entre pH 4.3 et 10.5. Ainsi, la stabilité du mutant Leu68Glu serait réversible lors de son arrivée dans l'abomasum, un avantage que ne possédait pas ses prédécesseurs.

### 3.3 Comparaison avec les précédents homologues de MB-1

La saga de la famille MB-1 se poursuit depuis 12 ans. Les stratégies utilisées pour améliorer sa stabilité et sa durée de vie *in vivo* ont été nombreuses et très souvent fructueuses. Le tableau 3.3 résume les améliorations de stabilité apportées à MB-1 avec les stratégies pour y arriver.

On peut y voir les différentes améliorations données par les différents mutants. Notre mutant Leu68Glu apparaît comme étant le meilleur mutant synthétisé jusqu'à maintenant, dépassant le MB-1 initial, la protéine utilisée pour les présentes mutations, MB-1Trp, les variants utilisant les ponts disulfures et des combinaisons de mutations. Ses caractéristiques le placent comme un mutant de choix pour l'expression hétérologue dans un organisme à intérêt agroalimentaire.

Les mutants portant la substitution Leu19Glu ont un problème de mauvais repliement qui expose leur tryptophane. Cette caractéristique a déjà été observée dans le cas des mutants MB-1TrpLH et MB-1TrpRH, où le pont disulfure intramoléculaire entre l'hélice 1 et 4 créait une tension dans la structure qui rendait le tryptophane accessible au solvant. Dans les deux cas, le mauvais comportement du cœur hydrophobe semble indiquer que ces protéines s'accumuleraient peu dans une plante transgénique.

Protéine	T <sub>m</sub> (°C)	Résistance protéolytique (% de protéine intacte après 60 minutes)	Stratégie	Référence
MB-1	39	0	Design d'une protéine nutritive	Beauregard et al. 1995
MB-1Cys	39	40	Dimérisation par pont disulfure intermoléculaire	Morrison et al. 2000
MB-1LH	48	43	Pont disulfure en topologie main gauche	Williams et al. 2002 Doucet et al. 2002
MB-1RH	50	35	Pont disulfure en topologie main droite	Williams et al. 2002 Doucet et al. 2002
MB-1Trp	55	60	Insertion d'un tryptophane dans le coeur	Gagnon et al. 2000
MB-1TrpLH	42	17	Combinaison de pont S-S main gauche avec le Trp	Doucet et al. 2002
MB-1TrpRH	49	40	Combinaison de pont S-S main droite avec le Trp	Doucet et al. 2002
Leu19Glu	47	-	Pont salin de surface avec le Trp	Sasseville et al. 2003*
Leu68Glu	61	65	Pont salin de surface avec le Trp	Sasseville et al. 2003*
Leu19Glu/Leu68Glu	43	35	Pont salin de surface avec le Trp	Sasseville et al. 2003*

\* Les résultats du présent travail.

**Figure 3.3: Compilation des stratégies appliquées à MB-1 depuis sa création et des stabilités thermodynamiques correspondantes.** Tableau des caractéristiques thermodynamiques des différents homologues de la protéine nutritive de design *de novo* MB-1. Les stratégies appliquées y sont également inscrites.



### 3.4 Perspectives de recherche

Malgré les améliorations apportées à sa stabilité conformationnelle, Leu68Glu pourrait encore être amélioré de plusieurs façons. Tout d'abord, sa dénaturation thermique n'est pas encore aussi coopérative que celle d'une protéine naturelle, ce qui semble indiquer que son cœur hydrophobe est encore relativement fluide. Des efforts d'optimisation de la spécificité d'empaquetage des chaînes latérales du cœur hydrophobe devraient aider à améliorer ses caractéristiques. Les stratégies rationnelles ayant leurs limites, surtout pour une protéine dont la structure tridimensionnelle n'est pas connue, l'approche par évolution accélérée semble une approche intéressante. Un remodelage du cœur hydrophobe par mutation aléatoire des positions prédites dans ce cœur pourrait donner à MB-1 une fluidité réduite et donc une résistance aux protéases augmentée.

La poursuite de l'optimisation des interactions électrostatiques sur MB-1 est justifiée. Ces interactions sont connues dans de nombreuses protéines naturelles comme étant des éléments de spécification de repliement natif. Le design de ponts salins directement dans le cœur hydrophobe, sans nécessairement être stabilisant, peut servir à contraindre une protéine à adopter une forme unique, comme c'est le cas pour lysozyme du phage T4 (Anderson et al. 1990). Sans être dans le cœur, ces interactions électrostatiques pourraient être ajoutées aux interfaces comme dans les cas décrits à la section 9. Ces substitutions aux interfaces pourraient se faire par design rationnel ou encore par une approche aléatoire de dégénérescence des positions d'interfaces, avec une

sélection à haut débit basée sur la stabilité thermique et la résistance aux protéases (par "phage display" ou "RNA display") (Chu et al. 2002; Riechmann et al. 2000).

Une autre approche intéressante en design rationnel serait de créer un site de liaison au métal en utilisant les coudes 1 et 3. Ce type d'approche a déjà démontré son efficacité pour améliorer les caractéristiques des protéines *de novo* (Hill et al. 2000). De plus, certains domaines de liaisons d'ions divalents contiennent une proportion particulièrement élevée de méthionine, dont les atomes de soufre servent à la complexation de l'ion (Wernimont et al. 2003). Donc, tout en stabilisant la protéine en créant un tel domaine, il serait possible d'augmenter le taux de méthionine. Il serait également possible de conserver le taux de méthionine en remplaçant des méthionines déstabilisantes dans la structure par celles participant au domaine.

D'un point de vue académique uniquement, une expérience de remplacement de toutes les méthionines du cœur hydrophobe nous donnerait peut-être une vision toute autre de la protéine. Peut-être que le design de MB-1 est parfait à la base et que la seule chose qui le rend difficile à domestiquer est la relative polarité du noyau hydrophobe. Celui-ci contient plusieurs chaînes latérales de méthionines dont les atomes de soufre possèdent deux paires d'électrons libres qui augmentent cette polarité. De plus, la grande flexibilité de la chaîne latérale de la méthionine lui permet un beaucoup plus grand nombre de conformères que d'autres chaînes latérales hydrophobes, ce qui nuit à l'obtention d'un état natif unique. Une protéine sans méthionine dans le cœur hydrophobe serait peut-être plus facile à cristalliser ou son patron de diffraction aurait peut-être une

plus haute résolution. La structure tridimensionnelle résolue pourrait être utilisée pour designer de nouveaux sites d'insertion de méthionine moins déstabilisants. Elle nous permettrait également de cibler les problèmes de repliement ou d'identifier les régions repliées d'une façon différente de celle prédite.

## BIBLIOGRAPHIE

-Alcocer MJ, Murtagh GJ, Bailey K, Dumoulin M, Meseguer AS, Parker MJ, Archer DB. 2002. The disulphide mapping, folding and characterisation of recombinant Ber e 1, an allergenic protein, and SFA8, two sulphur-rich 2S plant albumins. *J Mol Biol* 324:165-175.

-Altenbach SB. 1987. Cloning and sequence analysis of a cDNA encoding a Brazil nut protein exceptionally rich in methionine. *Plant Mol Biol* 8:239-250.

-Anderson DE, Becktel WJ, Dahlquist FW. 1990. pH-induced denaturation of proteins: a single salt bridge contributes 3-5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry* 29:2403-2408.

-Beach L, Tarczynski MC. 2000. Methods of increasing accumulation of essential amino acids in seeds. US Patent 6,127,600.

-Beauregard M, Goraj K, Goffin V, Heremans K, Goormaghtigh E, Ruyschaert JM, Martial JA. 1991. Spectroscopic investigation of structure in octarellin (a de novo protein designed to adopt the alpha/beta-barrel packing). *Protein Eng* 4:745-749.

-Beauregard M, Hefford M A, Teather RM. 1994. Detection of  $\beta$ -galactosidase  $\alpha$ -peptide fusions using a commercially available monoclonal antibody. *Biotechniques* 16:831-834, 836, 838.

-Beauregard M, Teather RM, Hefford MA. 1995. Design, expression and initial characterization of MB-1, a de novo protein enriched in essential amino acids. *Bio/Technology* 13:974-981.

-Betz SF, Liebman PA, DeGrado WF. 1997. De novo design of native proteins: characterization of proteins intended to fold into antiparallel, rop-like, four-helix bundles. *Biochemistry* 36:2450-2458.

-Boerjan W, Bauw G, Van Montagu M, Inze D. 1994. Distinct phenotypes generated by overexpression and suppression of S-adenosyl- l-methionine synthetase reveal developmental patterns of gene silencing in tobacco. *Plant Cell* 6:1401–1414.

-Brive L, Dolphin GT, Baltzer L. 1997. Structure and function of an aromatic ensemble that restricts the dynamics of the hydrophobic core of a designed helix-loop-helix dimer. *J Am Chem Soc* 119:8598-8607.

-Burkhard P, Ivaninskii S, Lustig A. 2002. Improving coiled-coil stability by optimizing ionic interactions. *J Mol Biol* 318:901-910.

-Chakraborty S, Chakraborty N, Datta A. 2000. Increased nutritive value of transgenic potato by expressing a nonallergenic seed albumin gene from *Amaranthus hypochondriacus*. *Proc Natl Acad Sci USA* 97:3724–3729.

-Chu R, Takei J, Knowlton JR, Andrykovitch M, Pei W, Kajava AV, Steinbach PJ, Ji X, Bai Y. 2002. Redesign of a four-helix bundle protein by phage display coupled with proteolysis and structural characterization by NMR and X-ray crystallography. *J Mol Biol* 323:253-262.

-Chui CF, Falco SC. 1995. A new methionine-rich seed storage protein from maize. *Plant Physiol* 107:291.

-Cordes MH, Davidson AR, Sauer RT. 1996. Sequence space, folding and protein design. *Curr Opin Struct Biol* 6:3-10.

-Dahiyat BI, Mayo SL. 1997b. De novo protein design: fully automated sequence selection. *Science* 278:82-87.

-Dahiyat BI, Sarisky CA, Mayo SL. 1997a. De novo protein design: towards fully automated sequence selection. *J Mol Biol* 273:789-796.

-de Alba E, Santoro J, Rico M, Jimenez MA. 1999. De novo design of a monomeric three-stranded antiparallel beta-sheet. *Protein Sci* 8:854-865.

-de Lumen BO, Galvez AF, Revilla MJ, Krenz DC. 1999. Molecular strategies to improve the nutritional quality of legume proteins. *Adv Exp Med Biol* 464:117-126.

-Doel MT, Eaton M, Cook EA, Lewis H, Patel T, Carey NH. 1980. The expression in *E. coli* of synthetic repeating polymeric genes coding for poly(L-aspartyl-L-phenylalanine). *Nucleic Acids Res* 8:4575-4592.

-Doucet A, Williams M, Gagnon MC, Sasseville M, Beauregard M. 2002. Engineering nutritious proteins: improvement of stability in the designer protein MB-1 via introduction of disulfide bridges. *J Agric Food Chem* 50:92-98.

-Falco SC, Guida T, Locke M, Mauvais J, Sandres C, Ward RT, Webber P. 1995. Transgenic canola and soybean seeds with increased lysine. *Bio/Technology* 13:577-582.

-Falco SC, Keeler SJ, Rice JA. 1996. Synthetic storage proteins with defined structure containing programmable levels of essential amino acids for improvement of the nutritional value of plants. US Patent 5,559,223.

-Forano E, Flint HJ. 2000. Genetically modified organisms: consequences for ruminant health and nutrition. *Ann Zootech* 49:255-271.

-Funahashi J, Takano K, Yamagata Y, Yutani K. 2000. Role of surface hydrophobic residues in the conformational stability of human lysozyme at three different positions. *Biochemistry* 39:14448-14456.

-Gagnon MC, Williams M, Doucet A, Beauregard M. 2000. Replacement of tyr62 by trp in the designer protein milk bundle-1 results in significant improvement of conformational stability. *FEBS Lett* 484:144-148.

-Galili G. 1995. Regulation of lysine and threonine synthesis. *Plant Cell* 7:899-906.

-Galili G, Hofgen R. 2002. Metabolic engineering of amino acids and storage proteins in plants. *Metab Eng* 4:3-11.

-Gassner NC, Baase WA, Mooers BH, Busam RD, Weaver LH, Lindstrom JD, Quillin ML, Matthews BW. 2003. Multiple methionine substitutions are tolerated in T4 lysozyme and have coupled effects on folding and stability. *Biophys Chem* 100:325-340.

-Grundy JE, Wirtanen LY, Beauregard M. 1998. Addition of a poly-(6X) His tag to Milk Bundle-1 and purification using immobilized metal-affinity chromatography. *Protein Expr Purif*. 13: 61-66.

-Gu H, Yi Q, Bray ST, Riddle DS, Shiau AK, Baker D. 1995. A phage display system for studying the sequence determinants of protein folding. *Protein Sci* 4:1108-1117.

-Guenoune D, Amir R, Badani H, Wolf S, Galili S. 2002. Combined expression of S-VSPalpha in two different organelles enhances its accumulation and total lysine production in leaves of transgenic tobacco plants. *J Exp Bot* 53:1867-1870.

-Harbury PB, Plecs JJ, Tidor B, Alber T, Kim PS. 1998. High-resolution protein design with backbone freedom. *Science* 282:1462-1467.

- Hefford MA. 1997. Engineering nutritious proteins. *Biotechnol Genet Eng Rev* 14:191-210.
- Heremans B, Jacobs M. 1995. Threonine accumulation in a mutant of *Arabidopsis-thaliana* (L.) Heynh with an altered Aspartate kinase. *J Plant Physiol* 146:249–257.
- Hill RB, DeGrado WF. 1998. Solution structure of  $\alpha_2D$ , a nativelike de novo designed protein. *J Am Chem Soc* 120:1138-1145.
- Hill RB, Raleigh DP, Lombardi A, DeGrado WF. 2000. De novo design of helical bundles as models for understanding protein folding and function. *Acc Chem Res* 33:745-754.
- Houbrechts A, Moreau B, Abagyan R, Mainfroid V, Preaux G, Lamproye A, Poncin A, Goormaghtigh E, Ruysschaert JM, Martial JA, et al. 1995. Second-generation octarellins: two new de novo (beta/alpha)<sub>8</sub> polypeptides designed for investigating the influence of beta-residue packing on the alpha/beta-barrel structure stability. *Protein Eng* 8:249-259.
- Huang XL, Catignani GL, Swaisgood HE. 1994. Relative structural stabilities of beta-lactoglobulins A and B as determined by proteolytic susceptibility and differential scanning calorimetry. *J Agric Food Chem* 42:1276-1280.
- Jaynes J, Langridge P, Anderson K, Bond C, Sands D, Newman C, Newman R. 1985. Construction and expression of sythetic DNA fragments coding for polypeptides with elevated levels of essential amino acids. *Appl Microbiol Biotechnol* 21:200-205.
- Jiang X, Bishop EJ, Farid RSJ. 1997. A de novo designed protein with properties that characterize natural hyperthermophilic proteins. *J Am Chem Soc* 119:838-839.



-Johansson JS, Gibney BR, Skalicky JJ, Wand AJ, Dutton PL. 1998. A native-like three- $\alpha$ -helix bundle protein from structure-based redesign: a novel maquette scaffold. *J Am Chem Soc* 120:3881-3886.

-Jung R, Falco SC. 2000. Transgenic corn with an improved amino acid composition. Paper presented at the 8th International Symposium on Plant Seeds. Gatersleben, Germany.

-Jung R, Hastings C, Coughlan S, Hu D. 1998. Alteration of amino acids compositions in seeds. U.S. Patent 5,850,016.

-Kamtekar S, Schiffer JM, Xiong H, Babik JM, Hecht MH. 1993. Protein design by binary patterning of polar and nonpolar amino acids. *Science* 262:1680-1685.

-Karchi H, Shaul O, Galili G. 1993. Seed specific expression of a bacterial desensitized aspartate kinase increases the production of seed threonine and methionine in transgenic tobacco. *Plant J* 3:721-727.

-Keeler SJ, Maloney CL, Webber PY, Patterson C, Hirata LT, Falco SC, Rice JA. 1997. Expression of de novo high-lysine alpha-helical coiled-coil proteins may significantly increase the accumulated levels of lysine in mature seeds of transgenic tobacco plants. *Plant Mol Biol* 34:15-29.

-Kim J, Leustek T. 2000. Repression of cystathionine gamma-synthase in *Arabidopsis thaliana* produces partial methionine auxotrophy and developmental abnormalities. *Plant Sci* 151:9-18.

-Koehl P, Levitt M. 1999. Structure-based conformational preferences of amino acids. *Proc Natl Acad Sci U S A* 96:12524-12529.

-Kohn WD, Hodges RS. 1998. De novo design of  $\alpha$ -helical coiled coils and bundles: models for the development of protein-design principles. *TIB Tech* 16:379-389.

-Kortt AA. 1991. Amino acid and cDNA sequence of a methionine-rich 2S protein from sunflower seed (*Helianthus annuus* L.). *Eur J Biochem* 195:329–334.

-Kumar S, Nussinov R. 2000. Fluctuations between stabilizing and destabilizing electrostatic contributions of ion pairs in conformers of the c-Myc-Max leucine zipper. *Proteins* 41:485-497

-Kumar S, Nussinov R. 2002. Close-range electrostatic interactions in proteins. *Chembiochem* 3:604-617.

-Lazar GA, Desjarlais JR, Handel TM. 1997. De novo design of the hydrophobic core of ubiquitin. *Protein Sci* 6:1167-1178.

-Lee KK, Fitch CA, Lecomte JT, Garcia-Moreno E B. 2002. Electrostatic effects in highly charged proteins: salt sensitivity of pKa values of histidines in staphylococcal nuclease. *Biochemistry* 41:5656-5667.

-Levitt M. 1976. A simplified representation of protein conformations for rapid simulation of protein folding. *J Mol Biol* 104:59-107.

-Liao HH. 1993. Thermostable mutants of kanamycin nucleotidyl-transferase are also more stable to proteinase-K, urea, detergents, and water-miscible organic solvents. *Enzyme Microb Technol* 115:286-292.

-Loladze VV, Ibarra-Molero B, Sanchez-Ruiz JM, Makhataдзе GI. 1999. Engineering a thermostable protein via optimization of charge-charge interactions on the protein surface. *Biochemistry* 38:16419-16423.

-Lounnas V, Wade RC. 1997. Exceptionally stable salt bridges in cytochrome P450cam have functional roles. *Biochemistry* 36:5402-5417.

-MacCallum J, Hefford MA, Omar S, Beauregard M. 1997. Prediction of folding stability and degradability of the de novo designed protein MB-1 in cow rumen. *Appl Biochem Biotechnol* 66:83–93.

-Malakauskas SM, Mayo SL. 1998. Design, structure and stability of a hyperthermophilic protein variant. *Nat Struct Biol* 5:470-475.

-Mazur B, Krebbers E, Tingey S. 1999. Gene discovery and product development for grain quality traits. *Science* 285:372–375.

-Meier M, Lustig A, Aebi U, Burkhard P. 2002. Removing an interhelical salt bridge abolishes coiled-coil formation in a de novo designed peptide. *J Struct Biol* 137:65-72.

-Merz A, Knochel T, Jansonius JN, Kirschner K. 1999. The hyperthermostable indoleglycerol phosphate synthase from *Thermotoga maritima* is destabilized by mutational disruption of two solvent-exposed salt bridges. *J Mol Biol* 288:753-763.

-Miller S, Janin J, Lesk AM, Chothia C. 1987. Interior and surface of monomeric proteins. *J Mol Biol* 196:641-656.

-Morrow JA, Hatters DM, Lu B, Hochtl P, Oberg KA, Rupp B, Weisgraber KH. 2002. Apolipoprotein E4 forms a molten globule. A potential basis for its association with disease. *J Biol Chem* 277:50380-50385.

-Muntz K, Christov V, Saalbach G, Saalbach I, Waddell D, Pickardt T, Schieder O, Wustenhagen T. 1998. Genetic engineering for high methionine grain legumes. *Nahrung* 42:125–127.

-Nishimura C, Dyson HJ, Wright PE. 2002. The apomyoglobin folding pathway revisited: structural heterogeneity in the kinetic burst phase intermediate. *J Mol Biol* 322:483-489.

-Nozaki Y, Tanford C. 1971. The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. Establishment of a hydrophobicity scale. *J Biol Chem* 246:2211-2217.

-Offredi F, Dubail F, Kischel P, Sarinski K, Stern AS, Van de Weerd C, Hoch JC, Prosperi C, Francois JM, Mayo SL, Martial JA. 2003. De novo backbone and sequence design of an idealized alpha/beta-barrel protein: evidence of stable tertiary structure. *J Mol Biol* 325:163-174.

-Ogihara NL, Weiss MS, Degrado WF, Eisenberg D. 1997. The crystal structure of the designed trimeric coiled coil coil-VaLd: implications for engineering crystals and supramolecular assemblies. *Protein Sci* 6:80-88.

-Papas A, Wu SH. 1997. Rumen-stable delivery systems. *Adv Drug Deliv Rev.* 28: 323-334.

-Parsell DA, Sauer RT. 1989. The structural stability of a protein is an important determinant of its proteolytic susceptibility in *Escherichia coli*. *J Biol Chem* 264:7590-7595.

-Permyakov EA, Berliner LJ. 2000. Alpha-Lactalbumin: structure and function. *FEBS Lett* 473:269-274.

-Pettigrew JE, Yang H. 1997. Protein nutrition of gestating sows. *J Anim Sci* 75:2723-2730.

-Phelan P, Gorfe AA, Jelesarov I, Marti DN, Warwicker J, Bosshard HR. 2002. Salt bridges destabilize a leucine zipper designed for maximized ion pairing between helices. *Biochemistry* 41:2998-3008.

-Pomerantz JL, Wolfe SA, Pabo CO. 1998. Structure-based design of a dimeric zinc finger protein. *Biochemistry* 37:965-970.

-Ponder JW, Richards FM. 1987. Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes. *J Mol Biol* 193:775-791.

-Poso D, Sessions RB, Lorch M, Clarke AR. 2000. Progressive stabilization of intermediate and transition states in protein folding reactions by introducing surface hydrophobic residues. *J Biol Chem* 275:35723-35726.

-Richards FM, Lim WA. 1993. An analysis of packing in the protein folding problem. *Q Rev Biophys* 26:423-498.

-Richardson JS, Richardson DC. 2002. Natural beta-sheet proteins use negative design to avoid edge-to-edge aggregation. *Proc Natl Acad Sci U S A* 99:2754-2759.

-Riddle DS, Santiago JV, Bray-Hall ST, Doshi N, Grantcharova VP, Yi Q, Baker D. 1997. Functional rapidly folding proteins from simplified amino acid sequences. *Nat Struct Biol* 4:805-809.

-Riechmann L, Winter G. 2000. Novel folded protein domains generated by combinatorial shuffling of polypeptide segments. *Proc Natl Acad Sci U S A* 97:10068-10073.

-Roesler KR, Rao AG. 2001. Rapid gastric fluid digestion and biochemical characterization of engineered proteins enriched in essential amino acids. *J Agric Food Chem* 49:3443-3451.

-Santos FA, Santos JE, Theurer CB, Huber JT. 1998. Effects of rumen-undegradable protein on dairy cow performance: a 12-year literature review. *J Dairy Sci* 81: 3182-3213.

-Schafmeister CE, LaPorte SL, Miercke LJ, Stroud RM. 1997. A designed four helix bundle protein with native-like structure. *Nat Struct Biol* 4:1039-1046.

-Schwehm JM, Kristyanne ES, Biggers CC, Stites WE. 1998. Stability effects of increasing the hydrophobicity of solvent-exposed side chains in staphylococcal nuclease. *Biochemistry* 37:6939-6948.

-Struthers MD, Cheng RP, Imperiali B. 1996. Design of a monomeric 23-residue polypeptide with defined tertiary structure. *Science*. 271:342-345.

-Tabe LM, Droux M. 2002. Limits to sulfur accumulation in transgenic lupin seeds expressing a foreign sulfur-rich protein. *Plant Physiol* 128:1137-1148.

-Takano K, Tsuchimori K, Yamagata Y, Yutani K. 2000. Contribution of salt bridges near the surface of a protein to the conformational stability. *Biochemistry* 39:12375-12381.

-Teather RM. 1985. Application of gene manipulation to rumen microflora. *Can J Anim Sci* 65:563-574.

-Tisi LC, Evans PA. 1995. Conserved structural features on protein surfaces: small exterior hydrophobic clusters. *J Mol Biol* 249:251-258.

- Virtanen AI. 1966. Milk production of cows on protein-free feeds. *Science* 153:1603–1608.
- Voet D, Voet JG. 1995. *Biochemistry*, Second Edition. New York: John Wiley & Sons. 1361 p. (p.764-776)
- Walsh ST, Sukharev VI, Betz SF, Vekshin NL, DeGrado WF. 2001. Hydrophobic core malleability of a de novo designed three-helix bundle protein. *J Mol Biol* 305:361-373.
- Wang LJ, Kong XD, Zhang HY, Wang XP, Zhang J. 2000. Enhancement of the activity of l-aspartase from *Escherichia coli* W by directed evolution. *Biochem Biophys Res Commun* 276:346-349.
- Wang W, Hecht MH. 2002. Rationally designed mutations convert de novo amyloid-like fibrils into monomeric beta-sheet proteins. *Proc Natl Acad Sci U S A* 99:2760-2765.
- Wernimont AK, Huffman DL, Finney LA, Demeler B, O'Halloran TV, Rosenzweig AC. 2003. Crystal structure and dimerization equilibria of PcoC, a methionine-rich copper resistance protein from *Escherichia coli*. *J Biol Inorg Chem* 8:185-194.
- Williams M, Gagnon MC, Doucet A, Beauregard M. 2002. Design of high essential amino acid proteins: two design strategies for improving protease resistance of the nutritious MB-1 protein. *J Biotechnol* 94:245-254.
- Wu LC, Kim PS. 1997. Hydrophobic sequence minimization of the alpha-lactalbumin molten globule. *Proc Natl Acad Sci U S A* 94:14314-14319.

-Xu G, Wang W, Groves JT, Hecht MH. 2001. Self-assembled monolayers from a designed combinatorial library of de novo beta-sheet proteins. *Proc Natl Acad Sci USA* 98:3652-3657.

-Yang W, Jones LM, Isley L, Ye Y, Lee HW, Wilkins A, Liu ZR, Hellinga HW, Malchow R, Ghazi M, Yang JJ. 2003. Rational design of a calcium-binding protein. *J Am Chem Soc* 125:6165-6171.